INDIVIDUAL VARIATION IN STRESS HORMONES AND BEHAVIORAL PROFILES REPRESENTED BY PERSONALITY AND PLASTICITY IN TREE SWALLOWS

A Thesis
by
KRISTEN RENEE CONTENT

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

December 2017
Department of Biology

INDIVIDUAL VARIATION IN STRESS HORMONES AND BEHAVIORAL PROFILES REPRESENTED BY PERSONALITY AND PLASTICITY IN TREE SWALLOWS

A Thesis
by
KRISTEN RENEE CONTENT
December 2017

APPROVED BY:

Lynn Siefferman, Ph.D.
Chairperson, Thesis Committee

Alex Bentz, Ph.D.
Member, Thesis Committee

Nathan Mowa, Ph.D.
Member, Thesis Committee

Zack Murrell, Ph.D.
Chairperson, Department of Biology

Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

INDIVIDUAL VARIATION IN STRESS HORMONES AND BEHAVIORAL PROFILES REPRESENTED BY PERSONALITY AND PLASTICITY IN TREE SWALLOWS

Kristen Renee Content
B.S., Appalachian State University
M.S., Appalachian State University

Chairperson: Lynn Siefferman

Behavioral and evolutionary ecologists have been increasingly interested in why individuals from the same population exhibit consistent differences in behavioral and physiological traits across contexts- a phenomenon referred to as 'animal personality.' Both the repeatability of traits and the extent to which the traits vary with the environment (between-individual plasticity) should be considered aspects of the animals' personality. However, only recently have researchers begun to discover the potential underlying factors that influence individual variation of behavioral and physiological traits. Because these behavioral traits have been shown to be heritable, related to fitness, and shaped by evolution, they may represent strategic behavioral types (i.e., stress-coping styles) that enable individuals to cope with their environment. Thus, how an individual's physiological system responds and adapts to ecologically relevant environmental stressors can affect not only their behavioral response but more importantly their fitness. In the present study, I used a reaction norm approach to assess if baseline and stress-induced corticosterone (CORT) and nest defense behavior are components of personality and plasticity ('E', 'I', 'I X E') and how

within- and between-individual variation of these components respond to ecologically-relevant environmental variation (weather, site), breeding stage (early versus late nestling rearing stages), and individual characteristics (e.g., age, body mass) in a wild population of adult female tree swallows (*Tachycineta bicolor*). Although baseline CORT was not repeatable, stress-induced CORT was moderately repeatability (r= 0.330; 'I') while nest defense behavior displayed high repeatability (r= 0.445; 'I'). Similarly, baseline CORT was very plastic while stress-induced CORT and nest defense behavior exhibited low to moderate plasticity. Personality was associated with the extent to which individuals adjust their phenotype as function of changing conditions (behavioral plasticity); individuals with higher stress-induced CORT had more plastic stress-induced CORT and more aggressive individuals had more plastic behavior. Together these results suggesting that behavioral and hormonal personality can constrain plasticity in those traits. Although I found no direct links between CORT and nest defense behavior, my study highlights the importance of repeatedly measuring individuals while they experience ecologically-relevant stressors. Future studies should investigate how such within- and between-individual variation in personality and plasticity influences fitness.

Acknowledgments

Most importantly, I would like to give immeasurable thanks to Aaron Fischer who has provided unimaginable emotional support throughout my time as a graduate student, especially when I was writing my thesis, and with him by my side, I was able to successfully reach my goal. I would also like to provide thanks to my parents, Tanya and Ron Content, who have always provided unlimited support and unfailing love. Through their actions of hard-work, determination, and perseverance, they are my ultimate inspiration in life, and because of them, I will never stop pursuing my dreams and passions. In addition to my parents, I would like to thank both grandparents for always cheering me on and encouraging me in graduate school. I am also eternally grateful for my sister, Kelly Content, who has emotional supported me and has always listened and never complained when I need someone to talk to. I would like to express my heartfelt gratitude to my advisor, Dr. Lynn Siefferman, for her hospitability, supportive speeches, and unlimited patience. Throughout my entire time as a graduate student, she has continuously provided feedback to the numerous drafts I have sent her despite how rough they were and has always challenged me to the fullest extent. I would also like to thank my other committee members, Dr. Alex Bentz and Dr. Chishimba Nathan Mowa, for always listening and answering my questions and providing support whenever I needed it. The CORT assays, statistically analyses, and understanding the complicated pathways of stress hormones would not have been completed or understood without their help, and I am immensely grateful for their time and effort. I am also thankful

for Tyler Pyle, Callie Younginer, Chloe Burdick, Suryaveer Singh, Rachel Meyers and numerous devoted volunteers and capstone students for their hard-work throughout the past three field seasons. Without their daily help, I would not have been able to obtain all of this data and achieve my goal. I am also extremely grateful for members in the Gangloff-Siefferman lab, whom have patiently listened to my never-ending rants and provided invaluable feedback. Lastly, this study would not have been obtainable without the financial support of grants from High Country Audubon Society, Sigma Xi, and the Biology department at Appalachian State University. Overall, this unforgettable experience has opened my eyes to the true world of hard-work and has shaped me into the person I am today.

Table of Contents

List of Tables

List of Figures

Foreword

Chapter 1 of this thesis will be submitted to *The Journal of Functional Ecology*, a bimonthly peer-reviewed journal published by Wiley-Blackwell on behalf of the British Ecological Society; it has been formatted according to the style guide for that journal.

Introduction

Recent studies from a wide variety of species suggest that individual animals express limited behavioral flexibility (i.e., plasticity) and often display consistent differences in patterns of behavior within and across contexts- a phenomenon referred to as 'animal personality' (and also 'coping styles,' 'behavioral types,' 'temperament,' or 'behavioral syndromes,' reviewed in Stamps & Groothuis 2010 and in van Oers & Naguib 2013). Some of the commonly measured behavioral traits include: exploration, aggression, boldness, and neophobia (reviewed in Sih, Bell & Johnson 2004 and in van Oers & Naguib 2013). More recently, researchers have applied the concept of animal personality to physiological traits such as hormonal stress response (reviewed in Cockrem 2013). Both stress hormones and behavioral traits have demonstrated significant repeatability and heritability and therefore can respond to selection (reviewed in Bell, Hankison & Laskowski 2009 and in Carere, Caramaschi & Fawcett 2010; Stedman et al. 2017). Further, there is evidence that stress hormone profiles may reflect or mediate variation in behavioral traits (reviewed in Carere, Caramaschi & Fawcett 2010) but also that environmental stressors can influence whether such covariation is apparent (reviewed in Killen et al. 2013 and in Taff & Vikousek 2016)

Researchers often investigate relationships between stress hormones and behavior by studying stress-coping styles (Koolhaas et al. 1999). These styles refer to different coping mechanisms that reflect suits of behavioral and physiological responses toward stressful conditions (Koolhaas et al. 1999; reviewed in Cockrem 2007 and in Cockrem 2013). For example, chickens from selection lines for high feather pecking behavior have low corticosterone and display more proactive personalities (bold-aggressive) while those from low pecking

behavior lines are reactive (shy-submissive) with high corticosterone response levels (Korte et al. 1997). However, stress-coping styles are typically studied at the population-level while fewer studies have focused on within-population relationships between stress hormones and behavioral traits (Fürtbauer et al. 2015). Further, studies of coping styles are often limited because researchers tend to measure individuals one time, yet incorporating a repeated measures design can yield valuable within-individual variance (Archard et al. 2012; Fürtbauer et al. 2015). Studies that repeatedly measure wild animals under ecologically-relevant environmental conditions should increase the understanding of both individual variation (reviewed in Westneat, Wright & Dingemanse 2015) and potential relationships between stress hormones and behavior (reviewed in Hau et al. 2016).

Phenotypic plasticity allows individuals to cope with unpredictable environmental conditions. Yet, to date, there is limited research focusing on how ecologically-relevant stressors impact within-individual variation in personality and thus accurately quantify phenotypic plasticity of physiological and behavioral personality traits (reviewed in Dingemanse et al. 2010). Until recently, many behavioral ecologists had the misconception that personality and plasticity are incompatible with one another (reviewed in Dingemanse et al. 2010; Carter, Goldizen & Heinsohn 2012), yet plasticity need not be completely absent if personality is present within the population (reviewed in Dingemanse et al. 2010; Dingemanse et al. 2012; Mathot et al. 2012). For example, when environmental conditions change, individuals can alter their physiological and behavioral response while maintaining consistent differences from one another (Mathot et al. 2012). Moreover, researchers have found covariance between personality and plasticity in behavioral studies (reviewed in Dingemanse et al. 2010; Mathot et al. 2012; Dingemanse & Wolf 2013). For example, mice that are more aggressive tend to be more plastic

in social environments (Natarajan et al. 2009). Tree swallows (*Tachycienta bicolor*) that respond more aggressively to human intrusion also display greater weather-mediated plasticity in behavior (Betini & Norris 2012). However, there are few studies that have investigated associations between personality and plasticity in stress hormones and, to date, no individual-level covariation has been found (reviewed in Hau et al. 2016).

Hormone signals often coordinate the process of altering phenotypes to meet challenging environments, and glucocorticoids (stress hormones) are the primary mediators of allostasis (Landys, Ramenofsky & Wingfield 2006). In birds, corticosterone (CORT) is the main glucocorticoid steroid hormone and is regulated by the hypothalamic-pituitary-adrenal axis (HPA) (Landys, Ramenofsky & Wingfield 2006; reviewed in Hau et al. 2016). The HPA mediates CORT levels, especially when there are environmental stressors present, such as predation, low food availability, competition, and unpredictable weather (Wingfield et al. 1998; Wingfield 2013). When individuals experience acute and challenging stressors, baseline CORT begins to increase within 3 minutes (Romero & Reed 2005), and reaches maximum stress-induced levels within 30-60 minutes (reviewed in Romero 2004). Baseline CORT maintains homeostasis which is imperative for survival (Goodman 2010). Acute environmental changes cause baseline CORT to produce multiple simultaneous physiological effects on the body including activation of vital cells (reviewed in Sapolsky, Romero & Munck 2000). These cells mediate the transcription of cellular products which help stress-induced CORT to send signals to initiate physiological responses that aid in survivorship like gluconeogenesis (the breakdown of protein, fat, and glycogens), that stop energy storage for lipids, amino acids, and glucose, and that increase cardiovascular abilities (reviewed in Sapolsky, Romero & Munck 2000). Moreover, stress-induced CORT is responsible for eliciting a response towards potentially acute life

threating stressors while ultimately aiding in the restoration of homeostasis (reviewed in Sapolsky, Romero & Munck 2000 and in Romero 2004).

Most stress-response studies measure animals during their reproductive season, and in this life history stage, both baseline and stress-induced CORT are typically higher in avian species (Romero 2002; reviewed in Hau et al. 2016). This trend is likely present because individuals break down energy reserves to allocate toward reproductive processes and because environmental stressors are often greater during the breeding season (Romero 2002; Wingfield 2013). For example, when breeding white-crowned sparrows (*Zonotrichia leucophrys pugetensis*) experience relatively low ambient temperatures, their baseline and stress-induced CORT levels increase (Wingfield, Moore & Farner 1983). Further, song sparrows (*Melospiza melodia*) breeding in high-predation environments have higher baseline and stress-induced CORT compared to those breeding in low-predation environments (Travers et al. 2010). In addition to these environmental stressors, stress-induced CORT can also stimulate foraging behaviors when individuals have low energy levels (reviewed in Sapolsky, Romero & Munck 2000; Hau et al. 2010). Thus, individuals experiencing acute and challenging environmental situations may elicit an HPA response, which could affect their behavioral response (reviewed in Carere, Caramaschi & Fawcett 2010). Identifying whether stress hormones are mediators of behaviors may help explain within- and between-individual behavioral variation during unpredictable environments (reviewed in Cockrem 2007; Baugh et al. 2014; reviewed in Hau et al. 2016). For example, a study of three-spined sticklebacks (*Gasterosteus aculeatus*) demonstrates that basal cortisol levels predict shyness behavior but this relationship is primarily driven by within-individual, rather than between-individual, variance (Fürtbauer et al. 2015). This implies that a relationship between hormones and behavior may be missed if individuals are

only measured once, and highlights the importance of incorporating measures of within-individual variance when measuring stress hormones and behavioral traits (Fürtbauer et al. 2015).

To quantify personality and within-individual plasticity of stress physiology and behavioral traits across challenging environments, ecologists have applied the reaction norm statistical approach (reviewed in Dingemanse et al. 2010). Reaction norms measure changes in the phenotypic traits of individuals across environmental gradients (reviewed in Nussey, Wilson & Brommer 2007) and allow researchers to quantify relationships between personality and plasticity. This approach necessitates that the phenotypic trait of individuals is repeatedly measured in, at least, two environments (reviewed in Nussey, Wilson & Brommer 2007). Further, researchers often quantify personality via individual repeatability (R; intraclass correlation coefficient), which is calculated by the proportion of behavioral or physiological variation that is derived from within-individual variance divided by the sum of between- and within-individual variance (Lessells & Boag 1987; Dingemanse et al. 2012). Personality is displayed if there is low within-individual variance compared to the higher between-individual variance measured at the population-level (reviewed in Bell, Hankison & Laskowski 2009 and in Hau et al. 2016). Repeatability can also be quantified by the representation of the 'individual (I)' in reaction norm analyses, which is individual variation of the intercept (Dingemanse et al. 2012). In addition to personality, the reaction norms also estimate how the environment influences the overall average phenotype of the population ('E'), and the interaction between individual variance and the environment, represented as the slope (within-individual plasticity; 'I X E') through advanced mixed models (reviewed in Nussey, Wilson & Brommer 2007 and in Dingemanse et al. 2010; Dingemanse et al. 2012). More specifically, the 'I X E' interactions

reflect between-individual differences in plasticity (Westneat et al. 2011). Covariation between personality and plasticity can be detected by relationships between the intercept ('I') and slope ('I X E'; reviewed in Dingemanse et al. 2010). Reaction norm studies often reveal 1) that individuals differ consistently from the average phenotype of the population, 2) between-individual plasticity, and that 3) individual variation can be influenced by the environment (reviewed in Dingemanse et al. 2010).

I studied wild-living breeding female tree swallows and used a reaction norm approach to investigate how ecologically-relevant environmental variation (weather), breeding stage (early versus late nestling rearing stages), and individual characteristics (e.g., age, body mass) potentially influence stress hormones and parental behaviors. Twice, I collected phenotypic data from the same individuals, including blood hormone levels (baseline and stress-induced CORT) and behavioral measures. Specifically, I aimed (i) to assess repeatability of stress hormone levels and behavioral responses (i.e. personality); (ii) to assess plasticity to changes in ecologically relevant environmental variation by investigating the significant effects of 'E', 'I', and 'I X E' on stress hormone levels and behavioral responses; (iii) to test whether baseline and stress-induced CORT are associated; and (iv) to test whether baseline and stress-induced CORT levels (within and between) predict phenotypic variation in parental behavior.

Methods

Study Species

Wild breeding tree swallows are ideal model species in behavioral ecology (Betini & Norris 2012) and one of the most studied passerines in North America (Winkler et al. 2011).

Tree swallows are small (~20g) migratory passerines that typically eat aerial insects (Winkler et al. 2011). Food availability can be affected by weather (i.e., temperature, precipitation; Taylor 1963) and is generally unpredictable during the nestling stage (Hussell & Quinney 1987; Dunn et al. 2011). Tree swallows are secondary cavity nesters that readily use nest boxes. They typically arrive on the breeding grounds in early April and lay eggs in mid-May. Tree swallows produce a single brood each season, and clutch sizes of 5 and 6 eggs are most common (Winkler et al. 2011). Females incubate the clutch alone, but both sexes provide food to the nestlings, which typically fledge at 18–22 days. Both parents defend their nest against intruders, and there is high individual variation in nest defense behavior (Winkler 1992; Betini & Norris 2012). Although tree swallow nest defense behavior is not related to clutch size (Hainstock et al. 2010) or certainty of paternity (Dunn, Lifjelf & Whittingham 2009), it has been shown to vary with weather conditions (e.g., temperature; Betini & Norris 2012). Tree swallows show animal personality in territorial aggressive behavior (i.e., low within-individual variation coupled with high between-individual variation; Rosvall 2008) and in nest defense aggression (Betini & Norris 2012). Defending offspring, however, can be high-risk for parents and result in injury, or even death. Thus, nest defense behavior is a likely behavior used in unpredictable and stressful environmental situations, and therefore, maybe affected by CORT responses based on the individual's personality and plasticity component.

General Field Methods

I studied breeding tree swallows from May to July of 2014, 2015, and 2016 in Watauga County, North Carolina. In 2014, I monitored active nests in Meat Camp, NC (58 active nests; 36.299 N, 81.676 W), and, in 2015 and 2016, I collected nest data from two sites, Meat Camp

and Valle Crucis, NC (2015: 90; 2016: 85 active nests; 36.209 N, 81.778 W). I recorded first egg date, clutch size, and predicted hatch date using the average incubation period (15 days), and the assumption that incubation commences on the day the pent-ultimate egg is laid (Winkler et al. 2011). I monitored boxes for hatching and then recorded hatch date, clutch size, brood size, and the number of offspring that successfully fledged each nest. Twice, I 1) captured females to collect blood hormone samples and 2) observed nest defense behavior towards mock predators. I obtained weather data from the nearest National Oceanic and Atmospheric Administration (NOAA) weather station (Boone, NC). I calculated the average temperature and total precipitation of the 14 days before the date of each capture and behavioral observation as well as the mean daily temperature data on the dates of capture and behavioral observations (following Westneat et al. 2011).

Adult Capture and Blood Collection

I captured females between 06:00-15:00 using nest box traps (Stutchbury & Robertson 1987). I captured each female twice 1) on the day before the eggs were predicted to hatch (±3 day; hereto referred to as 'early nestling stage' of the breeding stage) and 2) when nestlings were 10 days old (±3 day; hereto referred to as 'late nestling stage' of the breeding stage). Upon first capture, I fitted females with a United States Geological Survey (USGS) numbered band and a colored leg band, and measured body mass and length of wing and tail. I estimated age as either second-year (SY) or after-second-year (ASY) using past capture records or plumage coloration as SY tree swallows are typically brown while ASY are iridescent in body plumage (Hussell 1983; Bentz & Siefferman 2013). Because it can be difficult to determine sex and identity of

swallows by sight using plumage and bands, I marked females with colored Sharpie® markers on their white abdominal feathers for subsequent behavioral observations.

Upon capture, I took blood samples to measure 'baseline' and 'stress-induced' concentrations of CORT using the standardized protocol for capturing and handling wild animals (Romero, Ramenofsky & Wingfield 1997; reviewed in Romero 2004). Within 3 min of capture, I took the baseline blood sample (time 0) and then, after keeping the bird in a bag, I took another blood sample at 30 min (time 30; stress-induced). I drew blood (~100µl) from the alar wing vein with a 26.5-gauge needle, transferred the blood into a 1.5mL tube and kept the sample on ice until it was centrifuged. Afterward, I stored plasma at -80° F until processing time.

Corticosterone Assays

I measured plasma CORT using an enzyme immunoassay kit (ENZO Life Sciences, Cat. No. ADI 901-097; Donkey anti-Sheep IgG; assay sensitivity 27 pg/mL). This type of assay plate has been used before to measure and compare CORT of passerines (Ouyang et al. 2011; Baugh et al. 2014). Duplicates of individual samples were assayed along with 5 standards that represented the standard curve ranging from 32 to 20,000pg/mL, blanks, and positive controls. For each year separately, baseline and stress-induced CORT of the early and late nestling stage were randomly placed among the assay plates. Assay plates shook for 2 hrs such that the conjugate corticosterone and antibody could bind together appropriately. I then washed wells 3 times using 400μl of wash solution and then incubated them in darkness for 1 hr after adding 200μl of p-nitrophenyl phosphate (pNpp) to each well. I added stop solution and then read on Victor Multilabel plate reader at 405nm. I used a four-point logistic standard curve to compute CORT concentrations. In 2014, CORT was measured across 5 assays with a 2.01% inter-assay

variation, and the average intra-assay variation was 2.04%. In 2015, CORT was measured across 5 assays with a 3.14% inter-assay variation, and an average intra-assay variation of 1.89%. Finally, for 2016, CORT was measured across only 3 assays and had 1.19% inter-assay variation, and the average intra-assay variation was 1.94%.

Nest Defense Behavior Trials

I observed female nest defense aggression (behavioral trials) to mock predators 1 day after capture. Therefore, I performed behavioral trials when nestlings were 2 days old (± 3 day; 'early nestling stage') and again when nestlings were 11 days old (± 3 day; 'late nestling stage'). I positioned a model American crow (*Corvus brachyrhynchos*) 0.3-meters above the nest box and observed female behavior from 40 m. During each 5 min trial, I recorded the number of attacks for female, which is defined as a complete dive within a 10 m radius of the nest box.

Statistical Analysis

I analyzed CORT and behavior data using SPSS v. 24 (IBM Corp 2016) and R statistical software v. 3.4.0 (R Development Core Team 2016). All tests were set at a significance level of $p \leq 0.05$ and were two tailed. Using Shapiro-Wilks tests, I found that CORT variables (baseline and stress-induced) deviated significantly from normal distributions (all $p < 0.05$). Thus, I log 10 transformed CORT variables to ensure that data conformed to normality and then removed extreme outliers (>3x interquartile range). Sample sizes for statistical tests vary, because I occasionally failed to capture a female and sometimes failed to draw enough blood for CORT assays. Moreover, the behavioral variable (# of attacks) also significantly deviated from normal distributions ($p < 0.05$) based on the Shapiro-Wilks tests. However, various transformations (i.e.,

log and square-root) did not conform attack data to normality and thus was considered to have a non-Guassian error distribution. I also removed extreme outliers (>3x interquartile range) from the attack data that could potentially bias the analysis using a non-Guassian error distribution. For all analyses, all continuous and level two categorical predictor variables were mean centered. This allows for the grand intercept (estimate of the intercept) to account for the average environmental variable in the model, and the fixed effect parameters of both continuous and categorical variables can be compared for further analysis (Araya-Ajoy, Mathot & Dingemanse 2015). Variables that differ between individuals are grand mean centered, which is where individual means are calculated and mean centered based on the overall mean of the population. However, variables that differ within individuals are group mean centered (van de Pol & Wright 2009; Westneat et al. 2011). All variables were mean centered except for the Pearson's correlation test (Baugh et al. 2014).

I first considered baseline and stress-induced CORT as separate variables, because they are regulated differently by the HPA axis and thus, the effect on the individual is different and because they respond differently to mineralocorticoid (MR) and glucocorticoid (GR) receptors (Baugh et al. 2014; reviewed in Hau et al. 2016). To test the basic assumption that the capture restraint protocol lead to an increase in CORT, I initially used a univariate Linear Mixed Models (LMM), which was fitted by the package 'nlme' in R 3.4.0 (R Development Core Team 2016; Pinheiro et al. 2017). In this model, baseline to stress-induced CORT was the dependent variable, 'Individual ID' was the random effect, and breeding stage was the fixed effect. For this model, I used the function 'lme' to assess if CORT significantly increased from baseline (time 0) to stress-induced (time 30; data reported below) CORT. I also assessed the possibility of interdependence between baseline and stress-induced CORT (Aim 3), Romeo 2004; Baugh et al.

2014) by initially running Pearson correlations between baseline and stress-induced CORT variables. In the Pearson's correlation, I used CORT values that were combined for all years (2014, 2015, and 2016) and separated with year because of potential year effects. In this analysis, CORT values were analyzed separately in the early and late nestling stages. Then a LMM was used to account for confounding factors and potential interactions. In the LMM, stress-induced CORT was the dependent variable, 'Individual ID' was the random factor, and breeding stage and baseline CORT were the predictor variables. Interactions between baseline CORT and any environmental variables that influenced stress-induced CORT were taken out if non-significant ($p > 0.05$). In all mixed model analyses, 'Individual ID' was denoted as the random effect to account for pseudoreplication (van de Pol & Wright 2009).

LRT and Multicollinearity

In models of repeatability and reaction norms, I used a -2 likelihood ratio test (LRT) to investigate whether the fixed and random effects increased the model's fit. Based on the LRT, fixed and random effects that did significantly improved model fit were kept in the final model (West, Welch & Galecki 2007). Specifically, the test statistic for LRT shows the difference in the -2 log-likelihood scores between the models that do and do not have the specific effect (fixed and random), and then is tested against a Chi-square distribution. The degrees of freedom (df) used in the LRT come from the difference in the df in the two models being compared. When performing the LRT, restricted maximum likelihood (REML) was used for the comparison of random effects while maximum likelihood (ML) was used for the comparison of fixed effects if the random effect remained the same in both models (Pinheiro & Bates 2000; West, Welch & Galecki 2007). In addition to using LRT, I tested for multicollinearity between all relevant

independent predictor variables in all mixed models. A correlation matrix was performed with any potential predictor variables that could be used in the mixed model for both CORT and behavior variables. If correlation values where $r \geq 0.7$, I removed one of the redundant variables from analyses (Pinheiro & Bates 2000). Julian hatch date and Julian first egg date were highly correlated ($r = 0.779$), and thus Julian first egg date was removed from analyses. Brood size and number fledged were also correlated ($r = 0.700$) and thus number fledged was removed from analyses.

Repeatability of CORT and Nest Defense Behavior

I measured repeatability using R package "rptR" for CORT and behavior (Stoffel, Nakagawa & Schielzeth 2017). For CORT variables, I used this package to calculate the repeatability of individual baseline and stress-induced CORT levels separately, during the early and late nestling stages (Aim 1). I also measured repeatability of individual CORT for each year to investigate potential year effects. Following Nakagawa and Schielzeth (2010), I calculated repeatability of baseline CORT and stress-induced CORT using LMM-based repeatability with Gaussian error distribution. In these models, CORT variables were the dependent variable and 'Individual ID' was the random effect. Repeatability of CORT was carried out by extracting the within- and between-variance components from the LMM and fitted using restricted maximum likelihood (REML). Here, I report the repeatability coefficient (R), standard error (SD), 95% confidence interval (CI), and p-value (via LRT) of the LMM-based repeatability model. I estimated repeatability confidence intervals using parametric bootstrapping for 1000 iterations. Furthermore, the statistical significance that represents the repeatability of CORT was tested by

likelihood ratio test (LRT) and if the 95% CI did not overlap with zero, this also suggested statistical significance (Nakagawa & Schielzeth 2010).

I used the 'rptR' package in R software to calculate the generalized linear mixed model (GLMM) based repeatability from early to late nestling stage (Aim 1; Stoffel, Nakagawa & Schielzeth 2017). I also tested repeatability of individual behavioral levels separately for each year to account for year effects. Furthermore, because behavior had a non-Guassian error distribution, I computed link scale repeatability with an overdispersion factor and a Poisson distribution from a GLMM (overdispersed Poisson GLMM with loglink). For this model, I calculated confidence intervals representing the repeatability of behavior using 1000 iterations of non-parametric bootstrapping. This GLMM-based model had behavior as the dependent variable and 'Individual ID' as the random effect. Following guidelines from Nakagawa and Schielzeth (2010), I report the link scale repeatability coefficient R, standard error, 95% confidence interval (CI) and p-value (via LRT) of the GLMM-based model. The 95% CI were reported to suggest statistical significance if the two values did not overlap with zero and the $p \leq 0.05$, which was based on the likelihood ratio test (LRT; Nakagawa & Schielzeth 2010).

Reaction Norms of CORT and Nest Defense Behavior

I used reaction norm analyses to investigate consistent individual differences (personality) and between-individual variation (plasticity) in baseline CORT, stress-induced CORT, behavior (attacks) and whether environmental factors contributed to the between-individual variation in plasticity. For CORT variables, I fitted LMM models using the 'lm' and 'lme' function in the statistical R package 'nlme' (Pinheiro et al. 2017). However, for behavior, I fitted GLMM models using the R package 'glmmTMB' (Magnusson et al. 2017) because

behavior data had a zero-inflated Poisson distribution (Warton 2005). Furthermore, to investigate personality (Aim 1) and plasticity (Aim 2), I conducted model comparisons for both CORT variables and behavior to assess the possible significant effects from the environment ('E'), the individual ('I'), and the interaction between individual and environment ('I X E') using the likelihood ratio test (LRT), which compares two models based on the presence or absence of a fixed or random effect (following: Pinheiro & Bates 2000; reviewed in Nussey, Wilson & Brommer 2007 and in Dingemanse et al. 2010; Westneat et al. 2011; Carter, Goldizen & Heinsohn 2012). Finally, if the model converged, I tested for a significant correlation between the individual ('I'; personality) and the interaction between individual and environment ('I X E'; plasticity) using the extracted values from the LMM (following: Betini & Norris 2012; Carter, Goldizen & Heinsohn 2012)

To assess the potential significant effects of 'E', the environmental variable (breeding stage) were either added or removed from the model as the fixed effect and for this model, the random intercept and slope were not present (an average population-level change in CORT or behavior between breeding stage). In this model, I included 'E' as breeding stage (fixed effect) and used either CORT or behavior as the dependent variable. Furthermore, to assess the potential significant effects of 'I', I added or removed 'Individual ID' (random intercept) from the model and allowed 'E' to remain in both models as the fixed effect. This comparison of models evaluates the presence of consistent hormonal or behavioral differences across the breeding stage (personality; individual differences of intercept between breeding stage). Finally, to assess whether a significant effect was present for the 'I X E' interaction, I added or removed the random slope of 'E' while 'I' ('Individual ID') remained as the random intercept, and 'E' also remained as the fixed effect throughout both models (plasticity; individual differences of slope

between breeding stage). For this model, the 'E' fixed effect variables that potentially differed significantly within individuals between the breeding stage were added to the model as the random slope (such as female mass, chick age, capture date, and environmental factors (i.e. average daily temperature (upon capture date), average temperature, and total precipitation over the previous 14 days prior to capture). If there was a significant 'I X E' interaction and the phenotypic trait is statistically repeatable ('I'), I extracted the intercept and slope values from the LMM or GLMM using the best linear unbiased predictor (BLUP) and correlated them to test for a significant positive or negative covariation of personality and plasticity (following Pinheiro & Bates 2000; Betini & Norris 2012; Carter, Goldizen & Heinsohn 2012). This covariance was tested using a Kendal's rank correlation, which used the BLUP values obtained from the LMM or GLMM (following Betini & Norris 2012).

CORT as a Mediator of Nest Defense Behavior

To investigate whether CORT was a proximate mediator of behavior (Aim 4), I used the package 'glmmTMB' to account for zero inflation (Poisson distribution) of the behavior data. For this model, behavior (attacks) was the dependent variable, 'Individual ID' was the random effect, between- and within- baseline and stress-induced CORT were added into the GLMM model separately as predictors, and breeding stage remained in the model at all times as a fixed effect. To account for between- and within-individual variance, between represented the mean of the two baseline or stress-induced CORT values (grand mean-centered). Within represented the baseline or stress-induced CORT values per individual (group mean-centered; van de Pol & Wright 2009; Araya-Ajoy, Mathot & Dingemanse 2015).

Year, Site, and Age effects on CORT and Nest Defense Behavior

I used a LMM to investigate the effects of year, site, and age and potential interactions of breeding stage on CORT and nest defense behavior. Dependent variables were CORT or behavior variables, 'Individual ID' was the random effect, and the fixed effects were the environmental covariates, breeding stage, and any interactions with breeding stage. Interactions with breeding stage and the environmental covariates were removed if non-significant ($p > 0.05$).

Ethical Note

This study followed guidelines for the Care and Use of Animals for Research, Teaching, or Demonstrations at Appalachian State University (#16-16). This study was also conducted under the Institutional Animal Care and Use Committee (IACUC) under USGS master banding permit #23563 (Lynn Siefferman). Each bird was minimally handled to reduce stress and physical harm.

Results

Sample Sizes for Baseline and Stress-Induced CORT Physiology

In total, I monitored 223 tree swallow nests over 3 years but sample sizes for CORT samples and behavior vary because I failed to collect enough blood for some individuals and the nest may have failed before I could do a behavioral trial. First, I collected 176 baseline CORT blood samples in the early nestling stage (2014= 42, 2015= 61, 2016= 73) and 80 during the late nestling stage (2014= 16, 2015= 21, 2016= 43). For stress-induced CORT, I collected blood from 164 females for the early nestling stage (2014= 40, 2015= 57, 2016= 67) and 81 for the late nestling stage (2014= 15, 2015= 21, 2016= 45). I collected blood for baseline CORT for both

nestling stages for 71 individuals (2014= 13, 2015= 22, 2016= 36). I collected blood for stress-induced CORT for 66 individuals for both breeding stages (2014= 14, 2015= 16, 2016= 36). In the early breeding stage, I collected baseline and stress-induced CORT samples for 150 individuals (2014= 36, 2015= 53, 2016= 61), but during the late breeding stage, I collected baseline and stress-induced CORT samples for 67 individuals (2014= 11, 2015= 17, 2016= 39). Furthermore, I have behavioral data for 182 individuals during the early nestling stage (2014= 43, 2015= 61, 2016= 78) and 147 individuals for the late nestling stage (2014= 32, 2015= 59, 2016= 57). However, I only have repeated behavior for 115 individuals during the nestling stage (2014= 23, 2015= 41, 2016= 51).

The amount of time elapsed between capture and completion of blood collection was not significantly correlated with baseline CORT concentration during either breeding stage (early nestling: r= 0.038, n= 176, p= 0.617; late nestling: r= 0.072, n= 85, p= 0.515). Time of day of blood collection was also not significantly correlated with baseline CORT during either breeding stage (early nestling: r= 0.240, n= 46, p= 0.109; late nestling: r= -0.127, n= 45, p= 0.406), nor was it correlated with stress-induced CORT (early nestling: r= 0.095, n= 38, p= 0.569; late nestling: r= -0.091, n= 38, p= 0.586). Moreover, the date of capture did not significantly correlate with baseline CORT (early nestling: r= 0.004, n=176, p= 0.960; late nestling: r= -0.138, n=85, p= 0.209), nor with stress-induced CORT during either breeding stage (early nestling: r= -0.036, n=106, p= 0.718; late nestling: r= -0.098, n=81, p= 0.386).

Comparison of Baseline to Stress-Induced CORT

In the LMM, for all years combined, CORT levels increased from baseline to stress-induced (time 0 min to time 30 min) during both breeding stages (early nestling: F(1,149)=

461.293, n= 150, p < 0.001; late nestling: F(1,66)= 61.866, n= 66, p< 0.001), thus the capture restraint protocol induced elevated CORT. This increase in CORT was present within all years for early nestling stage (early nestling: 2014: F(1,35)= 109.120, n= 36, p< 0.001; 2015: F(1,52)= 130.792, n= 53, p< 0.001; 2016: F(1,60)= 249.378, n= 61, p< 0.001). However, during the late nestling stage, this increase was statistically significant in 2015 and 2016 but not 2014 (; 2015: F(1,16)= 24.197, n= 17, p< 0.001; 2016: F(1,38)= 40.078, n= 39, p< 0.001; 2014: F(1,10)= 2.895, n= 11, p= 0.120).

Changes in Baseline and Stress-Induced CORT across the Breeding Stage

Using a LMM, with all years combined, I found that baseline CORT of individual females increased significantly across the breeding stage (F(1,70)= 13.921, n= 71, p< 0.001). In both 2014 and 2016, individuals had significantly higher CORT in late nestling compared to early nestling stage (2014: F(1,12)= 23.254, n= 13, p< 0.001; 2016: F(1,35)= 6.740, n= 36, p= 0.011). However, in 2015, baseline CORT levels of individuals did not vary with breeding stage (2015: F(1,21)=0.640, n= 22, p= 0.433). Moreover, using all years combined, stress-induced CORT of individual females did not differ significantly across the breeding stage (F(1,65)= 0.003, n= 66, p= 0.966). Within each year, stress-induced CORT of individual females did not differ significantly across the breeding stage (2014: F(1,13)= 0.455, n= 14, p= 0.512; 2015: F(1,15)= 0.479, n= 16, p= 0.500; 2016: F(1,35)= 0.800, n= 36, p= 0.377).

Phenotypic Correlation between Baseline and Stress-Induced CORT

Using data from all years combined, within the early nestling stage, there was a significant correlation between baseline and stress-induced CORT (p= 0.013; Table 1; Fig. 1),

indicating individuals with high baseline CORT also had high stress-induced CORT. When split by year, this significant positive correlation remained in 2016 ($p= 0.013$; Table 1), but was not significant in 2014 or 2015 (2014: $p= 0.178$; 2015: $p= 0.613$; Table 1). Using data from all years combined, there was a significant positive relationship between baseline and stress-induced CORT within the late nestling stage ($p= 0.005$; Table 1; Fig. 2). Again, this relationship was driven primarily by 2016 (2016: $p= 0.012$; 2015: $p= 0.823$; 2014: $p= 0.423$; Table 1).

A LMM was used to further explore the relationship between baseline and stress-induced CORT. Baseline CORT was a significant positive predictor of stress-induced CORT ($p= 0.028$; Table 2). There were no significant interactions between baseline CORT and any environmental variables on stress-induced CORT, so all other variables were removed from the final model ($p> 0.05$).

Repeatability of CORT and Behavior

I used a LMM-based repeatability model to test for repeatability of both baseline and stress-induced CORT variables. When all three years were combined, I found that baseline CORT was not significantly repeatable within individual females across the breeding stages ($R= 0.033$, $p= 0.412$; Table 3; Fig. 3). Although the adjusted repeatability of baseline CORT was also not significant when controlling for breeding stage, the R value did increase ($R_{adj}= 0.116$). Next, I split the dataset by year and calculated the repeatability of baseline CORT with breeding stage as the fixed effect. In each year, baseline CORT remained not significantly repeatable (2014: $R< 0.001$, $p> 0.05$; 2015: $R= 0.202$, $p= 0.197$; 2016: $R< 0.001$, $p> 0.05$; Table 3).

When all years were combined, individuals showed significantly repeatable stress-induced CORT across breeding stage ($R= 0.330$, $p= 0.003$; Table 3; Fig. 4). Also, adjusted

repeatability of stress-induced CORT was significant when controlling for breeding stage and there was little difference in the two R values (R_{adj}= 0.323). I split the dataset by year and calculated the repeatability with breeding stage as the fixed effect. In 2014 and 2015, stress-induced CORT was not significantly repeatable across the breeding stage (2014: R= 0.247, p= 0.211; 2015: R= 0.071, p= 0.438; Table 3). However, in 2016, individuals were significantly repeatable (2016: R= 0.440, p= 0.003; Table 3).

I used a GLMM-based repeatability model to test for repeatability of nest defense aggression (# of attacks). With all three years combined, nest defense aggression was significantly repeatable between the two breeding stages (R= 0.445, p< 0.001; Table 3; Fig. 5). Moreover, adjusted repeatability of nest defense aggression when controlling for breeding stage did not differ from the non-adjusted repeatability (R_{adj}= 0.445). Next, because of potential year effects with repeatability, I split the dataset by year. In each of the years, nest defense aggression was significantly repeatable across the breeding stage (2014: R= 0.637, p< 0.001; 2015: R= 0.077, p< 0.001; 2016: R= 0.886, p< 0.001; Table 3).

Reaction Norm of Baseline CORT

For baseline CORT, there was an average population-level response to the change in breeding stage ('E'); baseline CORT significantly increased from early to late nestling stage (p= 0.001; Table 4; Fig. 6). Individuals did not exhibit consistent differences in baseline CORT across the breeding stage ('I'; p= 0.253; Table 4). Next, in the reaction norm, I tested whether the individual slopes of baseline CORT varied significantly with one another ('I X E'). Baseline CORT showed a significant 'I X E' interaction with breeding stage as the random slope, thus there was a marginally significant between-individual plasticity in response to the breeding stage

(p= 0.059; τ= 0.277; Table 4; Fig 9). This suggests individuals differed in their baseline CORT in response to the breeding stage. Other than breeding stage, the only environmental variable that converged in the model and had a significant 'I X E' with baseline CORT was average temperature over the previous 14 days (p< 0.001; Table 4; Fig. 7). Thus, individuals differed significantly in how temperature influenced changes in their baseline CORT across the breeding stage.

Reaction Norm of Stress-Induced CORT

For stress-induced CORT, there was not a significant average population change across the breeding stage (p> 0.05; Table 4; Fig. 6). However, individuals did exhibit consistent differences in hormonal plasma levels across the breeding stage ('I'; p= 0.004; Table 4). Individuals were tested to identify whether the individual slopes of stress-induced CORT varied significantly in the reaction norm ('I X E'). Stress-induced CORT showed a significant 'I X E' interaction with breeding stage, thus, stress-induced CORT exhibited between-individual plasticity in response to the breeding stage (p= 0.012; τ= 0.378; Table 4 and 5; Fig 9). There was also a significant positive covariance between intercept (stress-induced CORT) and slope (breeding stage), suggesting that individuals with higher stress-induced CORT adjusted their stress-induced CORT more in response to the breeding stage (Kendall's rank correlation: τ= 0.300, n= 66, p< 0.001; Table 5; Fig. 8). Only breeding stage showed a significant 'I X E' interaction for stress-induced CORT; all other environmental variables failed to converge in the LMM of the reaction norm.

Reaction Norm of Nest Defense Behavior

For nest defense behavior, there was not a significant average population-level change across the breeding stage ('E'; p= 0.078; Table 4; Fig. 6). However, individuals exhibited consistent behavioral differences across the breeding stage ('I'; p< 0.001; Table 4). To measure plasticity, individuals were tested to identify whether their slope varied significantly across the breeding stage ('I X E'). Nest defense behavior showed a significant 'I X E' interaction with breeding stage, thus, nest defense behavior exhibited between-individual plasticity in response to the breeding stage. This suggests that individuals differ in their nest defense behavior in response to breeding stage. Specifically, for this analysis between nest defense behavior and breeding stage ('I X E'), the unstructured reaction norm model was unable to converge, and thus, the variance component model was used, which only has one df difference compared to two (p< 0.001; Table 4; Fig. 6). Other than breeding stage, the only environmental variable that converged in the model and had a significant 'I X E' with nest defense behavior was total precipitation (14 days) (p< 0.001; τ= 0.09; Table 5). Thus, individuals differ significantly in how total precipitation influences changes in their nest defense behavior across the breeding stage. This interaction showed a positive covariance between the intercept (nest defense behavior) and slope (total precipitation 14 days) (Kendall's rank correlation: τ= 0.222, n= 115, p< 0.001; Table 5; Fig. 9), suggesting that individuals that had a high nest defense behavior also experienced a greater change in nest defense behavior with total precipitation of the preceding 14 days in response to the breeding stage.

Baseline and Stress-Induced CORT Predicting Nest Defense Behavior

In this model, mean and within baseline CORT were not significant predictors of nest defense behavior ($p > 0.05$; Table 6). Furthermore, for mean and within stress-induced CORT variables, all nest defense behavior models failed to converge when stress-induced CORT was either the fixed effect or the random slope and thus, not a predictor of nest defense behavior.

Year, Site, and Age Effects on Baseline CORT

In the LMM, when breeding stage was a fixed effect, year had a significant effect on baseline CORT ($F(df) = 10.328(69)$, $p < 0.001$; Table 7). There were no significant two way interactions between year, breeding stage, and all other environmental variables on baseline CORT (all $p > 0.05$). To investigate the year effect, post hoc analyses were completed between year and breeding stage. There was a significant year effect on early nestling-stage for baseline CORT ($F(df) = 6.662(70)$, $p = 0.002$; Table 7); post hoc analyses revealed that this relationship was driven by significantly higher baseline CORT in 2015 compared to 2016 ($p = 0.002$; 2014: mean= 0.511, n= 13; 2015: mean= 0.705, n= 22; 2016: mean= 0.244, n= 36). However, during the late nestling stage, there was no significant year effect on baseline CORT ($F(df) = 2.576(70)$, $p = 0.084$; Table 7). Furthermore, there was not a site or age effect on baseline CORT (Site: $F(df) = 0.249(69)$, $p=0.619$; Age: $F(df) = 3.634(68)$, $p = 0.061$; Table 7). Also, there was not a significant interaction of breeding stage and site on baseline CORT nor was there a significant interaction of breeding stage and age on baseline CORT ($p > 0.05$).

Year, Site, and Age Effects on Stress-Induced CORT

Using LMM with breeding stage as a fixed effect, year did not have a significant effect on stress-induced CORT ($F(df)= 0.002(64)$, $p= 0.962$; Table 7). There were no significant two way interactions between year, breeding stage and all other environmental variables on stress-induced CORT (all $p> 0.05$). Although there was not a site effect on stress-induced CORT, there was a significant interaction between breeding stage and site on stress-induced CORT. The significant site effect was only present during the late nestling stage ($F(df)= 11.035(64)$, $p= 0.002$; Table 7) and this effect showed that there was significantly higher stress-induced CORT for birds in Valle Crucis compared to Meat Camp. Finally, there was not an age effect ($F(df)= 0.190(64)$, $p= 0.664$; Table 7) nor was there a significant interaction between breeding stage and age on stress-induced CORT (all $p> 0.05$).

Year, Site, and Age Effects for Nest Defense Behavior

The GLMM model failed to converge when trying to depict if there was a significant effect of year and any interactions between breeding stage and year on nest defense behavior. However, there were two significant interactions with an environmental variable and breeding stage on nest defense behavior. The first interaction was breeding stage and average temperature (previous 14 days) ($z= 2.549$, $p= 0.011$; Table 7). During the early nestling stage, individuals attacked less when the average temperature was lower. However, this was a weak relationship ($r= 0.141$) with only 2.1% of the variance explained by this interaction. The second interaction was breeding stage and total precipitation (previous 14 days) ($z= -2.987$, $p= 0.003$; Table 7). During the early nestling stage, individuals attacked more when the total precipitation was higher but this relationship was weak ($r= 0.164$) and only explained 2.7% the variation in behavior.

Moreover, there was a significant site effect on nest defense behavior ($z = -2.124$, $p = 0.034$; Table 7). Overall, individuals attacked more often at Valle Crucis compared to Meat Camp ($p < 0.05$). The interaction between breeding stage and site was non-significant ($p > 0.05$). Finally, there was not a significant effect of age ($z = 0.930$, $p = 0.352$; Table 7) nor was there an interaction between breeding stage and age on behavior ($p > 0.05$).

Discussion

The reaction norm approach allowed me to quantify personality and plasticity of three phenotypic traits (baseline CORT, stress-induced CORT and nest defense aggression), to explore potential covariance between personality and plasticity of each trait, and to investigate relationships between these three traits. My capture restraint protocol led to CORT elevation in nearly all birds. Further, I found potential evidence of interdependence between baseline and stress-induced CORT; females with high baseline CORT also exhibited high stress-induced CORT. Because I studied wild breeding birds, I also investigated how ecologically-relevant environmental variation (weather) and breeding stage (early versus late nestling rearing stages) influence stress hormones and parental behaviors. Individual plasticity in baseline CORT was much greater than that of stress-induced CORT and nest defense behavior. On the population-level, adult females had higher baseline CORT when rearing older compared to younger nestlings. For both stress-induced CORT and nest defense behavior, however, there was no concurrent change in phenotype across the breeding stage on the population-level. Moreover, individuals did not exhibit repeatable baseline CORT, thus baseline CORT should not be considered a personality trait. Individuals showed repeatable stress-induced CORT and nest

defense behavior, demonstrating evidence of stress-induced CORT and nest defense behavior as personality traits. Moreover, between-individual plasticity ('I X E' interactions) in baseline CORT was much greater than that of stress-induced CORT and nest defense behavior. Baseline CORT varied with breeding stage and temperature. To a lesser extent, stress-induced CORT was affected by breeding stage while nest defense behavior was affected by breeding stage and precipitation. Finally, I found no evidence that either baseline or stress-induced CORT levels (within and between) predicted phenotypic variation in parental behavior.

Interdependence of CORT

The capture-restraint protocol I used caused individuals to increase CORT in all years except during the late nestling stage of 2014 (Romeo 2004; reviewed in Hau & Goymann 2015). The failure to find the expected trend during the late nestling stage in 2014 could have occurred if some of the birds were captured soon after a stressful event in the wild (Baugh et al. 2014) or because this sample size was too small to detect a true effect ($n= 11$, $p= 0.120$). Moreover, with all years combined, the same individuals that exhibited higher baseline CORT also had higher stress-induced CORT during the early and late nestling stage, which suggests interdependence that could be the result of shared mechanisms regulating baseline and stress-induced CORT (Baugh et al. 2014). For example, both CORT variables overlap in the activation of mineralocorticoid (MR) and glucocorticoid (GR) receptors (reviewed in Hau et al. 2016). These mechanisms could also be sensitive to the similar internal and external stimuli (Baugh et al. 2014). Alternatively, the interdependence of baseline and stress-induced CORT could result from capturing a female after a stressful event (Baugh et al. 2014) but this is unlikely as females were captured randomly and were often caught while brooding young on the nest. After I split the

dataset into years, this trend was only evident for 2016 and not 2014 or 2015 and thus could suggest the importance of larger sample sizes. Moreover, in 2016 there were significant phenotypic correlations between baseline and stress-induced CORT during both the early and late nestling stages, this trend may be driven by between-individual correlation rather than within-individual correlation (Baugh et al. 2014). It is possible that using mixed effect models to partition the within- and between-individual variance could reveal the variance type that is driving this potential interdependence trend (Dingemanse & Dochtermann 2013; Baugh et al. 2014). This would provide insight into whether genetic, maternal, or environmental factors are driving the trend in 2016 (reviewed in Hau et al. 2016).

Reaction Norms ('E'): Environmental Effects on Population-Level Hormonal and Behavioral Phenotypes

Reaction norm analyses revealed that, across the breeding stage ('E'), there was an overall increase in the average population-level baseline CORT. This likely occurred because parenting during the late nestling stage is costlier because nestling provisioning demand is higher (Bonier et al. 2009) and females are expected to have depleted their energy reserves (reviewed in Hau et al. 2016). Stressors (like decreases in blood glucose levels or changes in ambient weather) may be more likely to evoke a physiological response during this time (reviewed in Cockrem 2013 and in Hau et al. 2016). Moreover, from a natural selection perspective, nestlings that are older are more valuable to the parents (Trivers 1972) and thus females may adopt a strategy of sustaining potentially damaging physiological stress to ensure the survival of their young (Ouyang, Hau & Bonier 2011; reviewed in Hau et al. 2016). These results provide some support for the CORT-adaptation hypothesis, which suggests that elevated baseline CORT during the late

nestling period can be adaptive (Bonier et al. 2009). Indeed, a study of female tree swallows breeding in Ontario demonstrated that females that exhibited higher baseline CORT during the nestling stage (10-12 days old) compared to the incubation staged achieved higher fitness (Bonier et al. 2009). Although I did not measure baseline CORT during early incubation, the trend of increased baseline CORT with nestling age suggests that elevated baseline CORT facilitates appropriate allocation of resources toward reproductive demands and that doing so can have potential fitness benefits. Moreover, there were also year effects for baseline CORT that shed light on how environmental variation can influence reproductive investment. Individuals in 2015 did not increase their baseline CORT from the early to late nestling stages. Further, during the late nestling stage, individuals in 2015 had higher baseline CORT than individuals in 2016. Together, these year effects suggest that, in 2015, individuals were relatively stressed during the early nestling stage and may not have additional energy reserves to allocate towards reproductive investment. These year differences in baseline CORT may reflect differences in environmental conditions such as food availability, weather conditions, and predation levels.

For stress-induced CORT, however, the population-level analysis did not reveal a significant effect of breeding stage ('E'). It may be that the internal and external stimuli associated with the elevated cost of reproduction do not affect stress-induced CORT traits as much as baseline CORT or that individuals cannot afford the reproductive costs of elevated stress-induced CORT (reviewed in Hau et al. 2016). Mechanistically, stress-induced CORT responds to acute challenges that have high potential for death (i.e., the body enters the "emergency life history stage"; Wingfield et al. 1998). During this time, all energy reserves are redirected toward survival of the individual (reviewed in Sapolsky, Romero & Munck 2000). Typically, once an individual reaches a certain blood concentration of stress-induced CORT, the

negative feedback system is stimulated and the body avoids damaging side effects of stress-induced CORT, such as the inhibition of reproduction and decrease in immune function (reviewed in Hau et al. 2016). Although the population-level analysis did not reveal a significant effect of breeding stage on stress-induced CORT, breeding stage interacted with field site to influence stress-induced CORT. When rearing older nestlings, individuals at Valle Crucis exhibited higher stress-induced CORT compared to those breeding at Meat Camp. This effect might stem from environmental or population-level differences between Valle Crucis and Meat Camp- sites that are < 20 km apart. For example, breeding density is higher at Valle Crucis and thus females may encounter acute challenges more often (i.e., competition of nests or resources; predation; Bentz, Navara & Siefferman 2013). Together, the year and site effects on CORT highlight the importance of measuring ecologically-relevant environmental variables when measuring stress hormones.

Across the breeding stage ('E'), there was no directional change in average population-level parental defense aggression. I was surprised by this result because nest defense aggressiveness should increase as the nestlings become more valuable to the parents (Trivers 1972; Barash 1975; Tolonen & Korpimaki 1995). My data, however, are consistent with nest defense behavior of tree swallows breeding in Ontario that also did not report a population response associated with nestling age (Betini & Norris 2012; reviewed in Brommer 2013). Overall, population-level responses of behavior may be context dependent and also vary across populations (reviewed in Brommer 2013). Indeed, nest defense behavior varied with field site; individuals at Valle Crucis attacked more often than those at Meat Camp. Although this effect was present during both nestling stages, this pattern mirrors how site influenced stress-induced CORT and thus may be similarly affected by the environment. The higher breeding density of

tree swallows at Valle Crucis could cause females to engage more often in territorial defense aggression (Bentz, Navara & Siefferman 2013) or aid their neighbors more often in mobbing potential nest predators (Male, Jones & Robertson 2006) and either of these scenarios may select for more aggressive females at the Valle Crucis site.

Reaction Norms ('I'): The Individual Effects on Hormonal and Behavioral Phenotypes

Within the breeding stage, 3.33 % of the variation in baseline CORT and 33.0 % of the variation in stress-induced CORT can be attributed to between-individual differences. Further, baseline CORT was not repeatable while stress-induced was, thus baseline CORT should not be considered an individual personality trait while stress-induced CORT should (Carter, Goldizen & Heinsohn 2012). The extent to which CORT is repeatable in the avian literature is somewhat equivocal; in wild populations, repeatability has been reported as zero to moderate (0-0.5) for both baseline and stress-induced CORT (reviewed in Hau et al. 2016). Heritability of baseline and stress-induced CORT has been estimated for juvenile tree swallows using sib-sib relationships, and yielded low but significant heritability (baseline: h^2: 0.13-0.14; stress-induced h^2: 0.18; Stedman et al. 2017). Indeed, studies have suggested that heritability estimates of CORT levels may partly explain between-individual variation in repeatability of traits that can be attributed to genetic variance (Lessells & Boag 1987; reviewed in Hau et al. 2016, but see Dohm 2002). My results differ from wild breeding tree swallows sampled in Ontario that revealed repeatable baseline CORT and non-repeatable stress-induced CORT, which varied seasonally and across years (Ouyang et al. 2011). One difference between Ouyang et al.'s (2011) and my study is that I captured birds within 2 weeks while their birds were recaptured across 2 months, suggesting that repeatability of CORT may be dependent on the assessment period (Ouyang et al.

2011). My study, however, is consistent with others that have found stress-induced CORT to be more repeatable than baseline CORT during the breeding season (Small & Schoech 2015; reviewed in Hau et al. 2016). This may occur because stress-induced CORT responds to acute and emergent challenges rather than non-life threating situations (reviewed in Hau et al. 2016) while baseline CORT is more labile and, thus, less repeatable because it is more responsive to changes in the environment (e.g., weather and social context), internal body conditions, and the individual's energy reserve (reviewed in Hau & Goymann 2015).

Understanding repeatability can be difficult because the measurement value can be context-dependent (Carter, Goldizen & Heinsohn 2012; Fürtbauer et al. 2015), and lability in CORT due to environmental variation can mask true repeatability and thus estimates of heritability (reviewed in Hau & Goymann 2015, Hau et al. 2016 and in Taff & Vitousek 2016; Stedman et al. 2017). For example, consistency of individual traits could be more apparent if there was no directional change in the overall average population-level ('E') (reviewed in Dingemanse et al. 2010) of baseline CORT across the breeding season. Nonetheless, when I used adjusted repeatability to control for breeding stage, I found no significant change in the repeatability of baseline CORT. My results also differ from a study of captive great tits (*Parus major*) which demonstrated repeatability of both baseline CORT and stress-induced CORT (Cockrem & Silverin 2002). Environmental factors can be minimized in captivity and thus captive studies are expected to yield more repeatable CORT profiles (Romero, Reed & Wingfield 2000; Baugh et al. 2014; reviewed in Hau et al. 2016). Although studies of captive animals can be insightful, conducting studies in the wild is crucial to understanding how CORT varies with the environment (Baugh et al. 2014). Furthermore, low repeatability of CORT under field conditions (like those I found for baseline CORT), highlights the fact that studies that only

sample individuals once may only be describing variation due to difference in environmentally-mediated states and not describing the hormonal phenotype of the animal (reviewed in Hau at al. 2016).

Behavior is expected to be more repeatable than CORT (reviewed in Bell, Hankison & Laskowski 2009; Ouyang et al. 2011; reviewed in Hau et al. 2016) and indeed, nest defense behavior was highly repeatable across the breeding stage. My data are consistent with an Ontario population of tree swallows that showed repeatable nest defense aggression (towards humans; Betini & Norris 2012) and from a Pennsylvania population that exhibited repeatable territorial aggression (Rosvall 2008). Together, these studies suggest that tree swallows exhibit behavioral personality traits (Rosvall 2008; reviewed in Bell, Hankison & Laskowski 2009; Betini & Norris 2012). Although I found that individuals were highly repeatable in nest defense aggression, in retrospect, I suspect that my measure of parental nest defense could have been more detailed. As I only measured number of attacks to the mock predator, I had a high number of trials that were recorded as zero attacks. In addition to quantifying attacks, Betini and Norris (2012) used a more comprehensive methods of also recording vocalizations toward stimuli and nearest approach distance to predator. Although I statistically accounted for overdispersion, I may have missed some of the within- and between-individual variance in nest defense; a more sensitive measure of nest defense behavior should be used in future studies. Despite the possibility of missing some variation in aggression, it may also be adaptive for tree swallows to exhibit limited plasticity of nest defense behavior because plasticity can be costly (DeWitt 1998; reviewed in Nussey, Wilson & Brommer 2007). For instance, if challenging environmental conditions are coupled with the stressors associated with reproduction, this can deplete the individual's energy reserves quickly; thus, individuals may experience selection pressure for higher consistency in behavior

(Dall, Houston & McNamara 2004). Moreover, behavioral repeatability can be maintained through a variety of other mechanisms including fluctuating selection and disruptive selection (reviewed in Dingemanse & Reale 2005).

Reaction Norms ('I X E') and Covariance ('I' by 'I X E'): Individual Differences in How the Environment Effects Hormonal and Behavioral Phenotypes

The reaction norm approach allows for the investigation of between-individual variance of plasticity ('I X E') across the environment (breeding stage) and potential covariance between personality ('I') and the 'I X E' interaction. Although not always evident, studies indicate that an 'I X E' interaction is a characteristic of natural variation and thus individuals differ in their response to the environment (reviewed in Brommer 2013). Moreover, if the trait of interest is considered a personality trait (repeatable; 'I') and an 'I X E' interaction is established, it is possible to explore covariance between the intercept ('I'; personality) and slope ('I X E'; plasticity; reviewed in Dingemanse et al. 2010 and in Westneat, Wright & Dingemanse 2015). A covariance relationship demonstrates the overall pattern of individual variance and how individuals change over the environment, for example a strong positive covariance would describe a scenario in which individuals with a high expression of the trait also show a strong positive slope across the environment (reviewed in Brommer 2013). The lack of significant in the covariance trend is often represented by individual slopes overlapping (i.e., crossing over).

In this study, I found evidence of 'I X E' interactions for all three traits measured (baseline CORT, stress-induced CORT, and nest defense behavior), suggesting ample between-individual variance in plasticity, that is, individuals differed in their hormonal and behavioral plasticity across the breeding stage. However, the 'I X E' interaction for baseline CORT with

respect to breeding stage was only marginally significant (p= 0.059), and yet, this relationship may still be biological relevant. Additionally, there was a more pronounced 'I X E' interaction with temperature (14 days), demonstrating that individuals differ in how temperature (14 days prior) across the breeding stage influenced their baseline CORT. Many studies have found that temperature influences baseline CORT and individuals often respond differently depending on their physiological state (reviewed in Hau et al. 2016). Nonetheless, only a few studies have found a significant 'I X E' interaction for hormonal reaction norms (Lendvai et al. 2014; Martin & Liebl 2014; reviewed in Hau et al. 2016). In experimental food restriction study, Lendvai et al. (2014) found that the baseline CORT of individual house sparrows (*Passer domesticus*) differed in response to unpredictability of food ('I X E'). Because they found significant repeatability for baseline CORT ('I'), they also estimated the covariance between personality and plasticity but found no evidence of a covariance trend. However, in my study, baseline CORT is not considered a personality trait ('I'; not repeatable), and therefore, I did not investigate correlations between baseline CORT and breeding stage.

Because stress-induced CORT is considered a personality trait ('I'), and an ('I X E') interaction was evident across the breeding stage, I investigated covariation between stress-induced CORT (personality) and breeding stage (plasticity). Indeed, I found significant positive covariance (p< 0.001), suggesting individuals with higher CORT were more plastic in response to the breeding stage. It may be that individuals with higher stress-induced CORT also have more GR receptors or a higher affinity for GR receptors and thus more stress-induced CORT can be produced (reviewed in Hau et al. 2016). Individuals with lower stress-induced CORT may be constrained by their 'personality' and therefore, have lower availability of GR receptors (reviewed in Hau et al. 2016). All additional models of stress-induced CORT with other

environmental variables failed to converge, suggesting that the environmental variables did not explain enough variance to reveal 'I X E' interactions or potential covariance. It may be because stress-induced CORT exhibited little variation across the breeding stage with respect to environmental variables (i.e. weather) and / or because my sample sizes were too small to detect interactions (Westneat et al. 2011).

The was an 'I X E' interaction between nest defense behavior and breeding stage, which suggests that individuals differ in their nest defense behavior across breeding stage. When testing for the covariation between personality (nest defense behavior) and plasticity (breeding stage), however, the model did not converge, implying no covariance trend ('I' by 'I X E') potentially because of a lack of variance and / or low sample size (reviewed in Brommer 2013). Interestingly, I found a significant 'I X E' interaction and positive covariance between nest defense behavior and total precipitation (14 days) across the breeding stage, indicating that more aggressive individuals were more plastic in their response to precipitation. Perhaps the more aggressive individuals are more sensitive to weather conditions, which may influence food availability or foraging ability. My data are consistent with tree swallows in Ontario describing an 'I X E' interaction and a positive covariance between aggression and plasticity with respect to weather conditions, moreover in that population of tree swallows, the interaction between personality and plasticity influenced reproductive success (Betini & Norris 2012). However, my results should be interpreted with caution; many of my birds did not attack the mock predator, if the same individuals consistently failed to attack, this may over inflate the subset of individuals that have low nest defense aggression. Thus, the plasticity trend may have had a greater influence on individuals that displayed moderate to high levels of nest defense behavior.

CORT as a Mediator of Nest Defense Behavior

Although my study highlights the importance of measuring ecological relevant factors using a repeated measures design, I found no evidence of a direct link between baseline or stress-induced CORT and nest defense behavior across the breeding stage- either using a mean population approach (between) or an individual-level covariance approach (within). Although baseline CORT did not significantly predict nest defense behavior across the breeding stage, similar models using stress-induced CORT did not converge. The lack of convergence may imply that stress-induced CORT is not a predictor of nest defense behavior, the by-product of a low sample size, or there is not enough variation in stress-induced CORT. Perhaps a true relationship between CORT and nest defense behavior might have been revealed if individuals were measured under different environmentally stress-related conditions. Indeed, a study of sea bass (*Dicentrarchus labrax*) failed to reveal a relationship between hormones and risk-taking behavior when food abundance was moderate, but under low food availability, a positive relationship existed suggesting that the physiological state of the individual may reveal relationships that might otherwise be masked (Killen, Marras & McKenzie 2011). Alternatively, nest defense behavior may not be associated because baseline CORT was not repeatable. Although I measured each individual twice, there was a strong population-level effect of breeding stage on baseline CORT. Patterns of baseline CORT could have been clearer if I had also repeatedly measured individuals within each breeding stage (reviewed in Brommer 2013 and in Hau et al. 2016). In three-spined sticklebacks (*Gasterosteus aculeatus*) only within, not between, basal cortisol levels predicted shyness behavior, however, this relationship was only apparent during pre-predation threat conditions (i.e., low-risk) (Fürtbauer et al. 2015). Within-individual variation could be masking the trend and thus measuring individual traits twice under

two different environmentally relevant contexts may show important relationships between CORT and behavior that were otherwise missed (reviewed in Killen et al. 2013 and in Brommer 2013). Also, quantifying GR receptors, steroid binding globulins in the blood, and even identifying potential blockers for hormonal gene expressions may reveal how variation in hormones are mediated and thus how within- and between-individual variation in stress hormones may influence behavior (reviewed in Carere, Caramaschi & Fawcett 2010 and in Hau et al. 2016). Alternatively, CORT and behavioral phenotypes may be completely decoupled in tree swallows and thus no relationship would be revealed under any conditions.

Conclusions

Wild populations can provide valuable insight regarding how individuals' phenotypes differ in their response to environmental stressors (Betini & Norris 2012). My study highlights the importance of measuring personality and plasticity of hormonal and behavioral phenotypes under ecologically-relevant environmental conditions. Using a reaction norm approach, I found that stress-induced CORT and nest defense behavior are personality ('I') traits. Baseline CORT, on the other hand, was highly variable (i.e. not repeatable) and, at the population-level, increased with the breeding stage ('E'). Moreover, because I found that all three traits exhibited between-individual variation in plasticity ('I X E'), my data suggest individuals differed in their response to the breeding stage. Personality was associated with the extent to which individuals adjust their phenotype as function of changing conditions (behavioral plasticity); suggesting that personality can constrain plasticity. Finally, although I found no direct link between CORT and nest defense behavior, my data add to the limited research that has examined the effect of environmental stressors on the relationship between hormones and behavior (reviewed in Killen et al. 2013).

My results highlight the importance of quantifying environmental variation that could drive within-and between-individual variation and relationships between personality and plasticity. Studies that repeatedly measure individuals under these types of stressors could be key to understanding the relationships between physiology and behavior and why some personality and plasticity phenotypes (stress-coping styles) exhibit higher fitness than others.

References

Araya-Ajoy, Y.G., Mathot, K.J. & Dingemanse, N.J. (2015) An approach to estimate short-term, long-term and reaction norm repeatability. *Methods in Ecology and Evolution*, **12**, 1462-1473.

Archard, G.A., Earley, R.L., Hanninen, A.F. & Braithwaite, V.A. (2012) Correlated behaviour and stress physiology in fish exposed to different levels of predation pressure. *Functional Ecology*, **3**, 637-645.

Barash, D.P. (1975) Evolutionary aspects of parental behavior: distraction behavior of the alpine accentor. *The Wilson Bulletin*, **87**, 367-373.

Baugh, A.T., Oers, K.V., Dingemanse N.J. & Hau, M. (2014) Baseline and stress-induced glucocorticoid concentrations are not repeatable but covary within individual great tits (*Parus major*). *General and Comparative Endocrinology*, **208**, 154-163.

Bell, A.M., Hankison, S.J. & Laskowski, K.L. (2009) The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **4**, 771-783.

Bentz, A.B., Navara, K.J. & Siefferman, L. (2013) Phenotypic plasticity in response to breeding density in tree swallows: an adaptive maternal effect?. *Hormones and behavior*, **4**, 729-736.

Bentz, A.B. & Siefferman, L. (2013) Age-dependent relationships between coloration and reproduction in a species exhibiting delayed plumage maturation in females. *Journal of Avian Biology*, **44**, 1.

Betini, G.S. & Norris, D.R. (2012) The relationship between personality and plasticity in tree swallow aggression and the consequences for reproductive success. *Animal Behaviour*, **1**, 137-143.

Bonier, F., Moore, I.T., Martin, P.R. & Robertson, R.J. (2009) The relationship between fitness and baseline glucocorticoids in a passerine bird. *General and Comparative Endocrinology*, **1**, 208-213.

Brommer, J.E. (2013) Phenotypic plasticity of labile traits in the wild. *Current Zoology*, **4**, 485-505.

Carere, C., Caramaschi, D. & Fawcett, T.W. (2010) Covariation between personalities and individual differences in coping with stress: converging evidence and hypotheses. *Current Zoology*, **6**, 728-740.

Carter, A., Goldizen, A. & Heinsohn, R. (2012) Personality and plasticity: temporal behavioural reaction norms in a lizard, the Namibian rock agama. *Animal Behaviour*, **2**, 471-477.

Cockrem, J.F. (2007) Stress, corticosterone responses and avian personalities. *Journal of Ornithology*, **2**, 169-178.

Cockrem, J.F. (2013) Corticosterone responses and personality in birds: individual variation and the ability to cope with environmental changes due to climate change. *General and Comparative Endocrinology*, **190**, 156-163.

Cockrem, J.F. & Silverin, B. (2002) Variation within and between birds in corticosterone responses of great tits (*Parus major*). *General and Comparative Endocrinology*, **2**, 197-206.

Dall, S.R., Houston, A.I. & McNamara, J.M. (2004) The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecology Letters*, **8**, 734-739.

Dingemanse, N.J., Bouwman, K.M., van De Pol, M., Overveld, T.V, Patrick, S.C., Matthysen, E. & Quinn, J.L. (2012) Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *Journal of Animal Ecology*, **1**, 116-126.

Dingemanse, N.J. & Dochtermann, N.A. (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. *Journal of Animal Ecology*, **1**, 39-54.

Dingemanse, N.J., Kazem, A.J., Réale, D. & Wright, J. (2010) Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology & Evolution*, **2**, 81-89.

Dingemanse, N.J. & Réale, D. (2005) Natural selection and animal personality. *Behaviour*, **9-10**, 1159-1184.

Dingemanse, N.J. & Wolf, M. (2013) Between-individual differences in behavioural plasticity within populations: causes and consequences. *Animal Behaviour*, **5**, 1031-1039.

Dohm, M.R. (2002) Repeatability estimates do not always set an upper limit to heritability. *Functional Ecology*, **2**, 273-280.

Dunn, P.O., Winkler, D.W., Whittingham, L.A., Hannon, S.J. & Robertson, R.J. (2011) A test of the mismatch hypothesis: How is timing of reproduction related to food abundance in an aerial insectivore? *Ecology*, **2**, 450-461.

Dunn, P.O., Lifjeld, J. & Whittingham, L.A. (2009) Multiple paternity and offspring quality in tree swallows. *Behavioral Ecology and Sociobiology*, **63**, 911-922.

Fürtbauer, I., Pond, A., Heistermann, M. & King, A.J. (2015) Personality, plasticity and predation: linking endocrine and behavioural reaction norms in stickleback fish. *Functional Ecology*, **7**, 931-940.

Goodman, H.M. (2010) *Basic medical endocrinology*. Academic Press, Boston, MA.

Hainstock, M.H., Smith, M.C., Carr, J. & Shutler, D. (2010) Parental investment and brood value in tree swallows, *Tachycineta bicolor*. *Behaviour*, **4**, 441-464.

Hau, M., Casagrande, S., Ouyang, J.Q. & Baugh, A.T. (2016) Chapter two-glucocorticoid-mediated phenotypes in vertebrates: Multilevel variation and evolution. *Advances in the Study of Behavior*, **48**, 41-115.

Hau, M. & Goymann, W. (2015) Endocrine mechanisms, behavioral phenotypes and plasticity: known relationships and open questions. *Frontiers in Zoology*, **1**, S7.

Hau, M., Ricklefs, R.E., Wikelski, M., Lee, K.A. & Brawn, J.D. (2010) Corticosterone, testosterone and life-history strategies of birds. *Proceedings of the Royal Society of London B*, **1697**, 3203-3212.

Hussell, D.J. (1983) Age and plumage color in female tree swallows. *Journal of Field Ornithology*, **54**, 312-318.

Hussell, D.J. & Quinney, T.E. (1987) Food abundance and clutch size of tree swallows *Tachycineta bicolor*. *Ibis*, **S1**, 243-258.

IBM Corp (2016) IBM SPSS: Statistics for Windows, Version 24.0. Armonk, NY.

Killen, S.S., Marras, S. & McKenzie, D.J. (2011) Fuel, fasting, fear: routine metabolic rate and food deprivation exert synergistic effects on risk-taking in individual juvenile European sea bass. *Animal Ecology*, **80**, 1024–1033.

Killen, S.S., Marras, S., Metcalfe, N.B., McKenzie, D.J. & Domenici, P. (2013) Environmental stressors alter relationships between physiology and behaviour. *Trends in Ecology & Evolution*, **11**, 651-658.

Koolhaas, J.M., Korte, S.M., Boer, S.F.D., Vegt, B.J.V.D., Reenen, C.G.V., Hopster, H. *et al.* (1999) Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, **7**, 925-935.

Korte, S.M., Beuving, G., Ruesink, W.I.M. & Blokhuis, H.J. (1997) Plasma catecholamine and corticosterone levels during manual restraint in chicks from a high and low feather pecking line of laying hens. *Physiology & Behavior*, **3**, 437-441.

Landys, M.M., Ramenofsky, M. & Wingfield, J.C. (2006) Actions of glucocorticoids at a seasonal baseline as compared to stress-related levels in the regulation of periodic life processes. *General and Comparative Endocrinology*, **2**, 132-149.

Lessells, C.M. & Boag, P.T. (1987) Unrepeatable repeatabilities: a common mistake. *Auk*, **104**, 116-121.

Magnusson, A., Skaug, H.J., Nielsen, A., Berg, C.W., Kristensen, K., Maechler, M., van Bentham, K.J., Bolker, B.M. & Brooks, M.E. (2017) glmmTMB: Generalized linear mixed models using template model builder. R package version 0.1.4.

Male, S.K., Jones, J. & Robertson, R.J. (2006) Effects of nest-box density on the behavior of Tree Swallows during nest building. *Journal of Field Ornithology,* **1**, 61-66.

Martin, L.B. & Liebl, A.L. (2014) Physiological flexibility in an avian range expansion. *General and Comparative Endocrinology*, **206**, 227-234.

Mathot, K.J., Wright, J., Kempenaers, B. & Dingemanse, N.J. (2012) Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. *Oikos*, **7**, 1009-1020.

Nakagawa, S. & Schielzeth, H. (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, **4**, 935-956.

Natarajan, D., Vries, H.D., Saaltink, D.J., Boer, S.F.D. & Koolhaas, J.M. (2009) Delineation of violence from functional aggression in mice: an ethological approach. *Behavior Genetics*, **1**, 73-90.

Nussey, D.H., Wilson, A.J. & Brommer, J.E. (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, **3**, 831-844.

Ouyang, J.Q., Hau, M. & Bonier, F. (2011) Within seasons and among years: when are corticosterone levels repeatable? *Hormones and Behavior*, **5**, 559-564.

Ouyang, J.Q., Sharp, P.J., Dawson, A., Quetting, M. & Hau, M. (2011) Hormone levels predict individual differences in reproductive success in a passerine bird. *Proceedings of the Royal Society of London B*, **1717**, 2537-2545.

Pinheiro, J.C. & Bates, D.M. (2000) *Mixed-effects models in S and S-Plus*. Springer, New York.

Pinheiro, J.C., Bates D.M., DebRoy, S., Sarkar, D. & R Core Team (2017) nlme: Linear and nonlinear mixed effects models. R package version: 3.1-131.

R Development Core Team (2016) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rosvall, K.A. (2008) Sexual selection on aggressiveness in females: evidence from an experimental test with tree swallows. *Animal Behaviour*, **5**, 1603-1610.

Romero, L.M. (2002) Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *General and Comparative Endocrinology*, **1**, 1-24.

Romero, L.M. (2004) Physiological stress in ecology: lessons from biomedical research. *Trends in Ecology & Evolution*, **5**, 249-255.

Romero, L.M., Ramenofsky, M. & Wingfield, J.C. (1997) Season and migration alters the corticosterone response to capture and handling in an Arctic migrant, the white-crowned sparrow (*Zonotrichia leucophrys gambelii*). *Comparative Biochemistry and Physiology C*, **2**, 171-177.

Romero, L.M., Reed, J.M. & Wingfield, J.C. (2000) Effects of weather on corticosterone responses in wild free-living passerine birds. *General and Comparative Endocrinology*, **1**, 113-122.

Romero, L.M., & Reed, J.M. (2005). Collecting baseline corticosterone samples in the field: is under 3 min good enough? *Comparative Biochemistry and Physiology A*, **1**, 73-79.

Sapolsky, R.M., Romero, L.M. & Munck, A.U. (2000) How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews*, **1**, 55-89.

Sih, A., Bell, A. & Johnson, J.C. (2004) Behavioral syndromes: an ecological and evolutionary overview. *Trends in Ecology & Evolution*, **7**, 372-378.

Small, T.W. & Schoech, S.J. (2015) Sex differences in the long-term repeatability of the acute stress response in long-lived, free-living Florida scrub-jays (*Aphelocoma coerulescens*). *Journal of Comparative Physiology B*, **1**, 119-133.

Stamps, J. & Groothuis, T.G. (2010) The development of animal personality: relevance, concepts and perspectives. *Biological Reviews*, **2**, 301-325.

Stedman, J.M., Hallinger, K.K., Winkler, D.W. & Vitousek, M.N. (2017) Heritable variation in circulating glucocorticoids and endocrine flexibility in a free-living songbird. *Journal of Evolutionary Biology*, **30**, 1724-1735.

Stoffel, M.A., Nakagawa, S. & Schielzeth, H. (2017) rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. R package version 0.9.2.

Stutchbury, B.J. & Robertson, R.J. (1987) Do nest building and first egg dates reflect settlement patterns of females? *Condor*, **89**, 587-593.

Taff, C.C. & Vitousek, M.N. (2016) Endocrine flexibility: optimizing phenotypes in a dynamic world? *Trends in Ecology & Evolution*, **6**, 476-488.

Taylor, L.R. (1963) Analysis of the effect of temperature on insects in flight. *The Journal of Animal Ecology*, **32**, 99-117.

Tolonen, P. & Korpimaki, E. (1995) Parental effort of kestrels (*Falco tinnunculus*) in nest defense: effects of laying time, brood size, and varying survival prospects of offspring. *Behavioral Ecology*, **4**, 435-441.

Travers, M., Clinchy, M., Zanette, L., Boonstra, R. & Williams, T. D. (2010) Indirect predator effects on clutch size and the cost of egg production. *Ecology Letters*, **8**, 980-988.

Trivers, R.L. (1974) Parent-offspring conflict. *Integrative and Comparative Biology*, **14**, 249-264.

van Oers, K. & Naguib, M. (2013). Avian personality. *Animal Personalities- Behavior, Physiology, and Evolution* (eds C. Carere & D. Maestripieri), pp. 66-85. University of Chicago Press, Chicago, IL.

van de Pol, M. & Wright, J. (2009) A simple method for distinguishing within-versus between-subject effects using mixed models. *Animal Behaviour*, **3**, 753-758.

Warton, D.I. (2005) Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics*, **3**, 275-289.

West, B.T., Welch, K.B. & Galecki, A.T. (2007) *Linear mixed models: a practical guide using statistical software*. Chapman and Hall, Boca Raton, Florida.

Westneat, D.F., Hatch, M.I., Wetzel, D.P. & Ensminger, A.L. (2011) Individual variation in parental care reaction norms: integration of personality and plasticity. *The American Naturalist*, **5**, 652-667.

Westneat, D.F., Wright, J. & Dingemanse, N.J. (2015) The biology hidden inside residual within-individual phenotypic variation. *Biological Reviews*, **3**, 729-743.

Wingfield, J.C., Maney, D.L., Breuner, C.W., Jacobs, J.D., Lynn, S., Ramenofsky, M. & Richardson, R.D. (1998) Ecological bases of hormone—behavior interactions: the "emergency life history stage." *American Zoologist*, **1**, 191-206.

Wingfield, J.C., Moore, M.C. & Farner, D.S. (1983) Endocrine responses to inclement weather in naturally breeding populations of white-crowned sparrows (*Zonotrichia leucophrys pugetensis*). *Auk*, **100**, 56-62.

Wingfield, J.C. (2013) Ecological processes and the ecology of stress: the impacts of abiotic environmental factors. *Functional Ecology*, **1**, 37-44.

Winkler, D.W. (1992) Causes and consequences of variation in parental defense behavior by Tree Swallows. *Condor*, **94**, 502-520.

Winkler, D.W., Hallinger, K.K., Ardia, D.R., Robertson, R.J., Stutchbury, B.J. & Cohen, R.R. (2011) Tree swallow (*Tachycineta bicolor*). *The birds of North America online* (ed. A. Poole). Cornell Lab of Ornithology, Ithaca, NY. http://bna.birds.cornell.edu/bna/species/011 doi: 10.2173/bna.11.

Tables

Table 1. Pearson's correlations between baseline and stress-induced corticosterone (CORT) of female tree swallows with all years and separated by year (2014, 2015, and 2016) all breeding stages.

Variable (Breeding stage and Year)	**Correlation**	***N***	***P*-value**
Breeding Stage: Early nestling			
All years	0.202	150	0.013
2014	0.229	36	0.178
2015	-0.071	53	0.613
2016	0.317	61	0.013
Breeding Stage: Late nestling			
All years	0.340	66	0.005
2014	0.233	14	0.423
2015	0.061	16	0.823
2016	0.397	39	0.012

Table 2. Linear mixed model results of baseline corticosterone (CORT) and breeding stage (early versus late nestling stage) predict stress-induced CORT of female tree swallows.

Dependent Variable	**Predictor Variable**	**Estimate ±SE**	***N***	***t*-value**	***P*-value**
Stress-induced CORT	Breeding Stage	-0.059 ± 0.055	51	-1.085	0.283
	Baseline CORT	0.142 ± 0.063	51	2.260	0.028

Table 3. Repeatability values of corticosterone (CORT) and nest defense behavior across the breeding stage of female tree swallows. Repeatability values represent both all years combined and data separated by year (2014, 2015, and 2016). Adjusted repeatability results have breeding stage as the fixed effect.

Variable	**Repeatability (R)**	**SE**	**Adjusted Repeatability**	**95 % CI**	***N***	***P*-value**
Baseline CORT						
All years	0.033	0.082	0.116	[0, 0.276]	71	0.412
2014	< 0.001	0.165	0.371	[0, 0.551]	13	> 0.05
2015	0.202	0.165	0.195	[0, 0.577]	22	0.197
2016	< 0.001	< 0.001	< 0.001	[0, < 0.001]	36	> 0.05
Stress-induced CORT						
All years	0.330	0.110	0.323	[0.088, 0.531]	66	0.003
2014	0.247	0.207	0.229	[0, 0.676]	14	0.211
2015	0.071	0.157	0.055	[0, 0.522]	16	0.438
2016	0.440	0.132	0.438	[0.150, 0.666]	36	0.003
Nest Defense Behavior						
All years	0.445	0.099	0.445	[0.303, 0.698]	115	< 0.001
2014	0.637	0.218	0.614	[0.254, 0.984]	23	< 0.001
2015	0.077	0.213	0.077	[0.004, 0.802]	41	< 0.001
2016	0.886	0.106	0.883	[0.580, 0.959]	51	< 0.001

Table 4. Reaction norm analyses for corticosterone (CORT) and nest defense behavior in female tree swallows: comparisons of models with and without fixed and random effects using log-likelihood ratio tests (LRT). Individual ID ('ID') is the random effect.

Comparison	Model	Log-Likelihood	*df*	*N*	Models Compared	X^2	*P*-value
	Baseline CORT						
E	M1: breeding stage as fixed effect	255.992	3	71	M1 vs M2	9.06	0.001
	M2: without breeding stage	265.0606	2				
I	M3: ID as random intercept	255.0368	4	71	M3 vs M4	0.95	0.253
	M4: without ID	255.992	3				
I x E	M5: ID as random intercept; breeding stage as random slope	250.7774	6	71	M5 vs M6	4.26	0.059
	M6: without breeding stage as random slope	255.0368	4				
	Stress-induced CORT						
E	M1: breeding stage as fixed effect	74.21956	3	66	M1 vs M2	-3.96	> 0.05
	M2: without breeding stage	78.21957	2				
I	M3: ID as random intercept	71.00622	4	66	M3 vs M4	7.18	0.004
	M4: without ID	78.18174	3				
I x E	M5: ID as random intercept; breeding stage as random slope	63.57402	6	66	M5 vs M6	7.43	0.012
	M6: without breeding stage as random slope	71.00622	4				
	Nest defense behavior						
E	M1: breeding stage as fixed effect	1639.2	3	115	M1 vs M2	2.4	0.078

Comparison	Model	Log-Likelihood	*df*	*N*	Models Compared	X^2	*P*-value
	M2: without breeding stage	1641.6	2				
I	M3: ID as random intercept	1209.8	4	115	M3 vs M4	429.4	< 0.001
	M4: without ID	1639.2	3				
I x E	M5: ID as random intercept; breeding stage as random slope	1070.2	5	115	M5 vs M6	139.6	< 0.001
	M6: without breeding stage as random slope	1209.8	4				

Table 5. Estimated variances for random slope and covariance of slope and intercept for stress-induced corticosterone (CORT) and nest defense behavior in female tree swallows in response to significant environmental variables. τ and p-value come from Kendall's rank correlation test.

Slope term	**Variance in slope (±SE)**	**Cov (± SE)**	τ	N	X^2	***P*-value**
Stress-induced CORT						
Breeding stage	-0.001 ± 0.045	0.112 ± 0.335	0.300	66	7.43	< 0.001
Nest defense behavior						
Precipitation (14 days)	0.198 ± 0.267	1.72 ± 1.31	0.222	115	132.50	< 0.001

Table 6. Effects of corticosterone (CORT) variables on nest defense behavior in female tree swallows. Using a generalized linear mixed model (GLMM), breeding stage was the fixed effect and Individual ID was the random effect. Between- and within-individual CORT variables were mean centered. The dash symbols indicate that the model failed to converge.

Dependent Variable	**Predictor Variable**	**Estimate (± SE)**	**N**	***z* value**	***P*-value**
Nest Defense Behavior					
	Mean baseline CORT	-0.356 ± 0.599	56	-0.593	0.553
	Within-baseline CORT	-0.153 ± 0.165	56	-0.930	0.353
	Mean stress-induced CORT	--	--	--	--
	Within stress-induced CORT	--	--	--	--

Table 7. Covariates (year, site, age and environmental variables) that could have effected corticosterone (CORT) and nest defense behavior of female tree swallows. Only significant interactions with breeding stage were added. Year did not converge in the nest defense behavior model. F-value is for CORT variables and z-values are for nest defense behavior.

Variable	**Effect ± SE**	***N***	**F or *z*-value**	***P*-value**
Baseline CORT				
Breeding Stage*Year	-0.205 ± 0.064	71	10.33	< 0.001
Site	-0.099 ± 0.199	71	0.249	0.619
Age	-0.442 ± 0.232	71	3.634	0.061
Stress-induced CORT				
Year	0.002 ± 0.039	66	0.002	0.962
Site	-0.061 ± 0.065	66	0.884	0.351
Breeding Stage*Site	-0.285 ± 0.086	66	11.04	0.002
Age	-0.034 ± 0.078	66	0.190	0.664
Nest Defense Behavior				
Breeding Stage*Temp (14days)	0.765 ± 0.300	115	2.55	0.012
Breeding Stage* Precip (14days)	-1.871 ± 0.627	115	-2.99	0.002
Site	-0.866 ± 0.408	115	-2.12	0.034
Age	0.442 ± 0.442	115	0.930	0.352

Figures

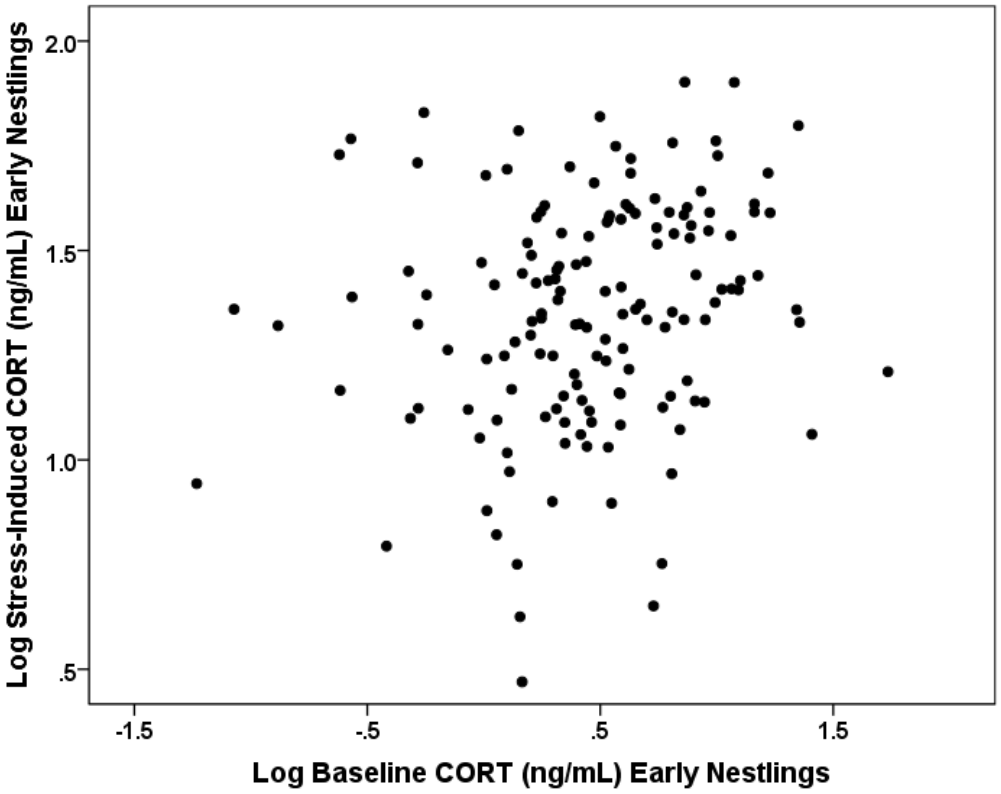

Fig. 1. Relationship between baseline and stress-induced corticosterone levels measured during the early (0+/- 3 days of hatching) nestling stage in a population of breeding tree swallows.

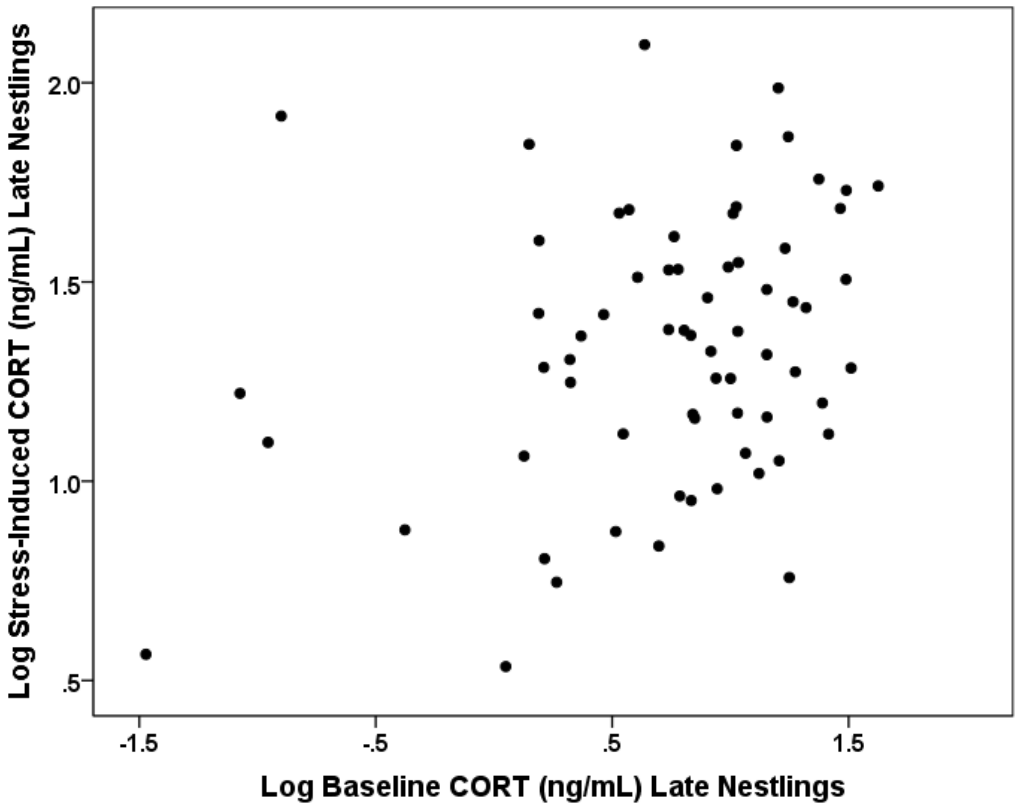

Fig. 2. Relationship between baseline and stress-induced corticosterone (CORT) levels measured during the late (10 +/- 3 days of hatching) nestling stage in a population of breeding tree swallows.

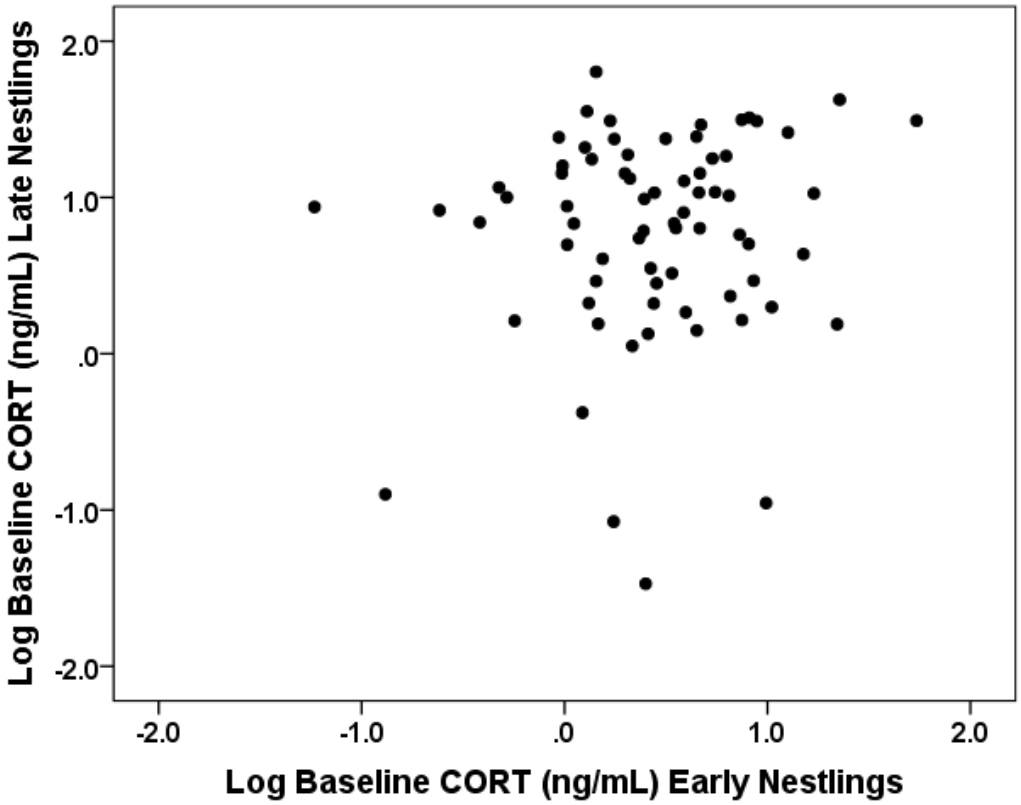

Fig. 3. Relationship between repeated measures of baseline corticosterone (CORT) (time: 0 min) measured during the early (0+/- 3 days of hatching) and late nestling stage (nestling age: 10 +/- 3 days) in a population of breeding tree swallows.

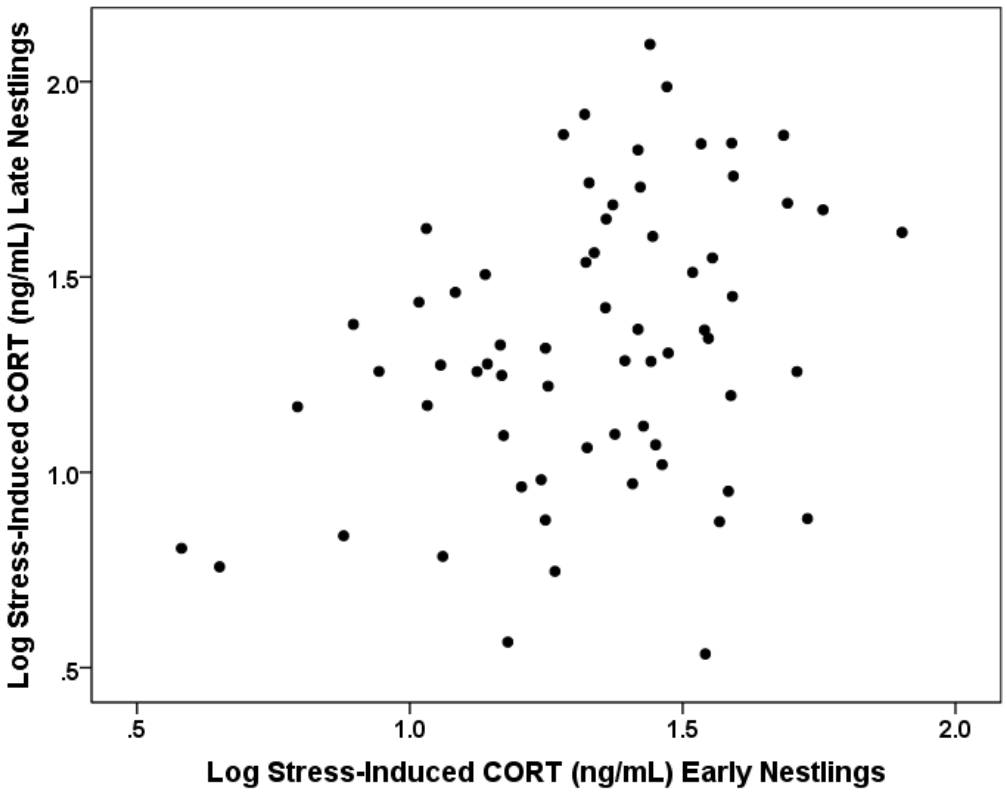

Fig. 4. Relationship between repeated measures of stress-induced corticosterone (CORT) levels (time: 30 min) measured during the early (0 +/- 3 days of hatching) and late nestling stage (nestling age: 10 +/- 3 days) in a population of breeding tree swallows.

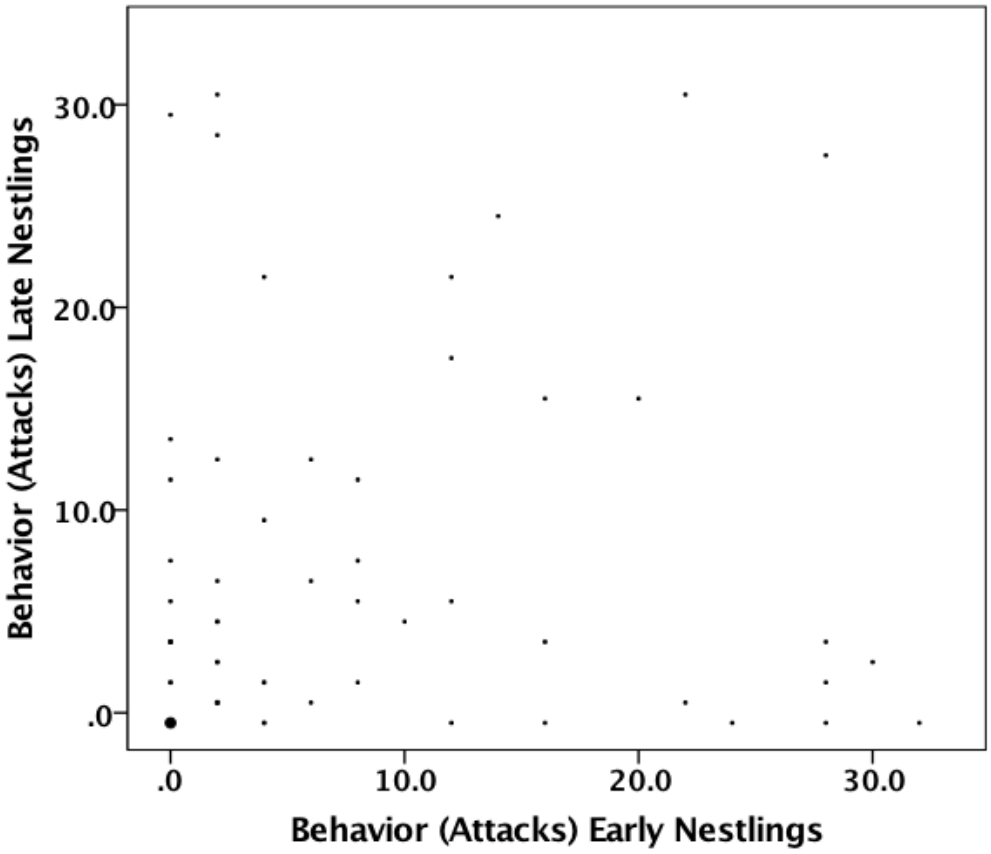

Fig. 5. Relationship between repeated measures of nest defense behavior measured during the early (0 +/- 3 days of hatching) and late nestling stage (nestling age: 10 +/- 3 days) in a population of breeding tree swallows. The size of data points reflect sample size (the large data point represents 50 birds).

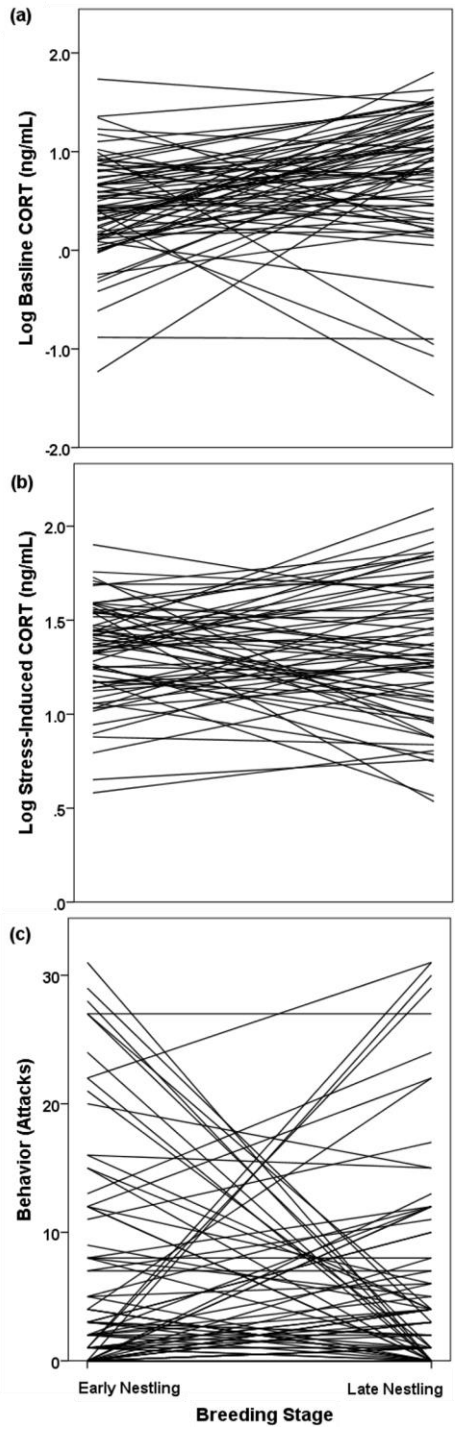

Fig. 6. Hormonal and behavioral individual plasticity in adult female tree swallows for (a) baseline corticosterone (CORT), (b) stress-induced corticosterone (CORT), and (c) nest defense behavior across the breeding stage. The X axis represents the early and late nestling stage of the breeding season and the Y axis is the observed hormonal and behavioral trait. Each line represents the raw or logged value of an individual's response to the breeding stage.

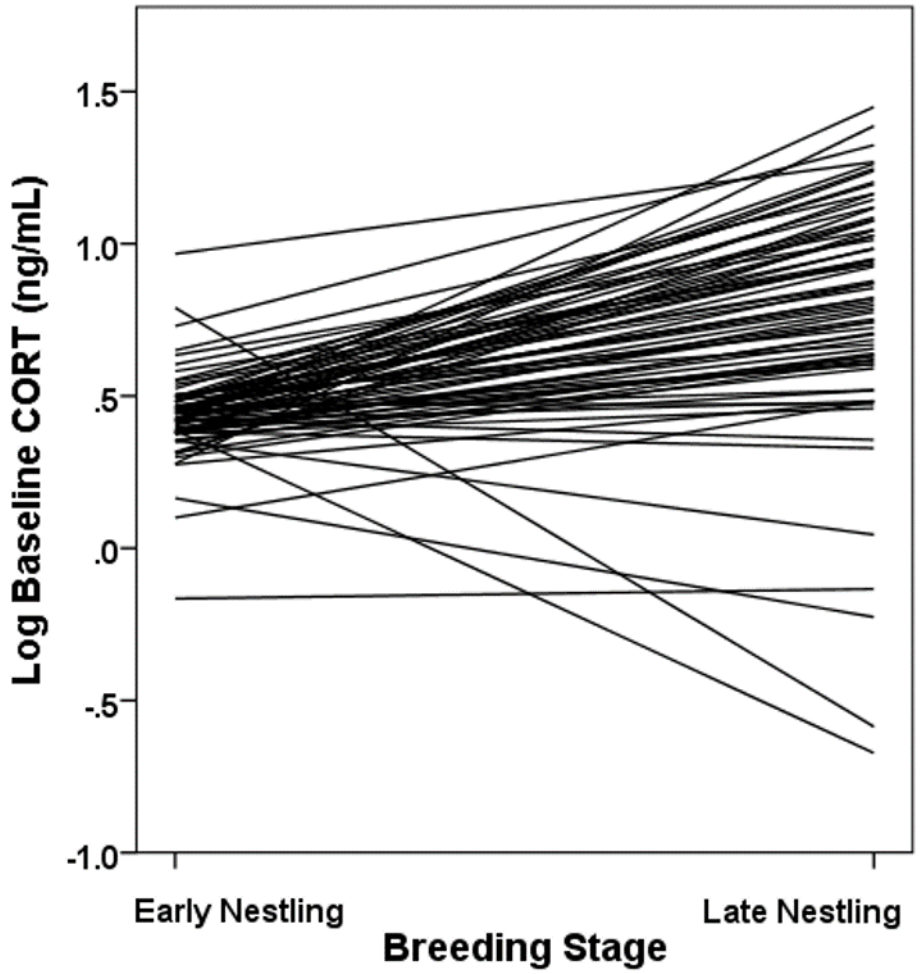

Fig. 7. Individual plasticity of baseline corticosterone (CORT) with respect to temperature (previous 14 days) across the breeding stage. The X axis represents the early and late nestling stage of the breeding season and the Y axis is the log of baseline corticosterone (CORT). Each line represent an individual's response and are statistically adjusted to temperature (previous 14 days) across the breeding stage (BLUP values).

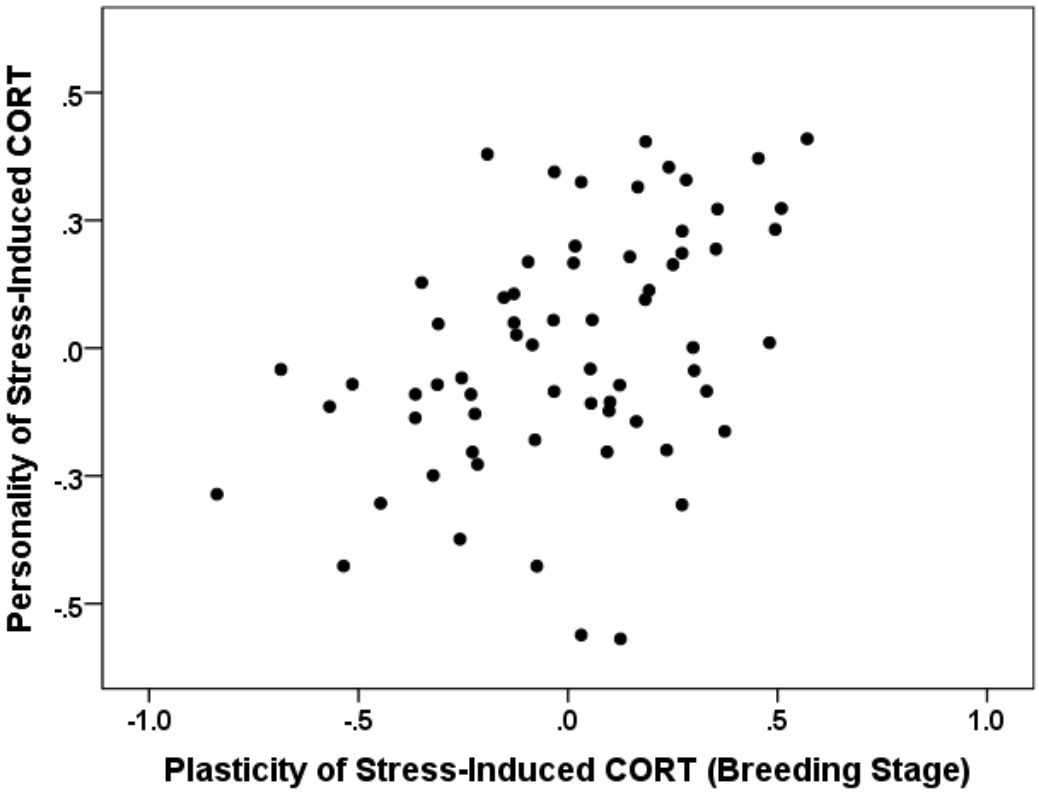

Fig. 8. Relationship between personality (intercept) and plasticity (slope) of stress-induced corticosterone (CORT) with respect to breeding stage of female tree swallows. Personality and plasticity values of stress-induced CORT are based on the random effect BLUP values extracted from the LMM.

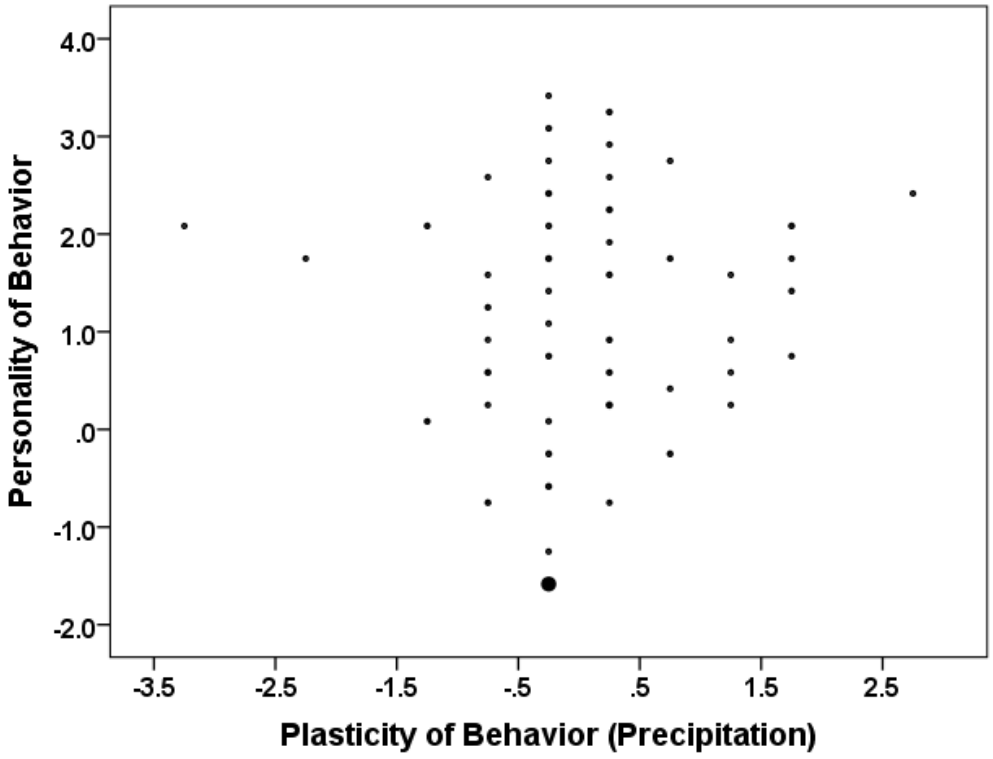

Fig. 9. Relationship between personality (intercept) and plasticity (slope) of nest defense behavior with respect to total precipitation (previous 14 days). Personality and plasticity values of nest defense behavior are based on the random effect BLUP values extracted from the GLMM. The size of data points reflect sample size (the large data point represents 40 birds).

Vita

Kristen R. Content was born in Columbus, Ohio in 1993. She attended elementary and middle school in Delaware County, OH and then moved half-way through high school to Cary, NC. She enrolled into college at Appalachian State University and started helping out in the Gangloff-Siefferman lab in 2011. She completed an undergraduate honors thesis and received a Bachelor of Science degree in Cell/Molecular Biology with a minor in Chemistry in May 2015. In fall 2015, she began her Master of Science degree at Appalachian State University in Biology where she pursued her passion for animals and biology. She received her Master of Science degree in Biology in December of 2017.